

UNIVERSIDADE DE LISBOA  
FACULDADE DE PSICOLOGIA



**ATUALIZAÇÃO ASSIMÉTRICA DE CRENÇAS  
ACERCA DO COMPORTAMENTO MORAL DOS  
OUTROS**

**Ana Maria de Oliveira Aparício Martins**

**MESTRADO INTEGRADO EM PSICOLOGIA**

**Área de Especialização em Cognição Social Aplicada**

**2020**

UNIVERSIDADE DE LISBOA  
FACULDADE DE PSICOLOGIA



**ATUALIZAÇÃO ASSIMÉTRICA DE CRENÇAS  
ACERCA DO COMPORTAMENTO MORAL DOS  
OUTROS**

**Ana Maria de Oliveira Aparício Martins**

**Dissertação orientada pelo Professor Doutor André Mata**

**MESTRADO INTEGRADO EM PSICOLOGIA**

**Área de Especialização em Cognição Social Aplicada**

**2020**

## **Agradecimentos**

Nos próximos parágrafos, gostaria de agradecer a todos aqueles que estiveram comigo ao longo deste percurso e que o tornaram possível. Agradeço,

Ao meu orientador, Professor Doutor André Mata, por me acolher no seu projeto e partilhar comigo as suas ideias. Pela disponibilidade, compreensão e apoio constantes. Por me ter transmitido segurança no decorrer de um processo que se revelou mais moroso e atribulado do que o previsto. Pelo seu interesse sincero. Por ser um exemplo de dedicação e entusiasmo contagiante pela investigação.

Aos meus colegas André Vaz e Rogério Gouveia e à Cristina Mendonça, pelos valiosos contributos em diferentes fases deste percurso, desde a troca de ideias inicial à análise de resultados, passando pela programação do estudo. Este trabalho não teria chegado a bom porto sem a sua colaboração.

Aos meus pais, pelo amor, pelo apoio, pela presença. Por terem acreditado em mim, sobretudo quando eu própria não consegui fazê-lo. Sei que não foi um período fácil, mas, à vossa maneira, nunca deixaram de ser casa, o lugar de conforto que tanto precisei. E ao meu irmão, sempre atento, sempre preocupado, cada vez mais um verdadeiro amigo.

À Inês, à Mariana e à Marta, pela paciência ilimitada, pelo apoio contínuo, pelos conselhos, pela motivação. Por me compreenderem. Por não me deixarem cair. Não tenho dúvidas que sem a vossa amizade, que levo comigo para sempre, este capítulo da minha história teria um desfecho muito diferente e o trabalho que aqui apresento não teria sequer sido concluído. Ter-vos a meu lado durante os últimos anos e, em particular, nesta fase crucial foi mais importante do que alguma vez conseguirei expressar por palavras. Obrigada, do fundo do coração.

Aos meus restantes familiares e amigos, por se importarem genuinamente e por me proporcionarem os melhores momentos de pausa e diversão, sem os quais a concentração e o empenho necessários teriam sido muito mais difíceis de alcançar. De modo particular, ao meu priminho Álvaro, que com o seu sorriso de bebé ajudou a revigorar as minhas forças em mais do que uma ocasião.

À Dr.<sup>a</sup> Ana Moniz, minha psicoterapeuta, pelo trabalho incansável que tem feito comigo. Pela disponibilidade e preocupação. Por ter entrado nesta viagem e me ter guiado

numa fase particularmente dura. Por ter sido um pilar indispensável. Tive muita sorte em tê-la encontrado a si.

A todos os participantes do estudo, por terem dispensado um pouco do seu tempo para contribuir para este projeto.

## Resumo

Apesar de vivermos, em grande parte, rodeados de pessoas que presumivelmente consideramos boas do ponto de vista moral, tendemos a pensar negativamente sobre o comportamento moral da humanidade. De acordo com a literatura em Psicologia Social a respeito dos vieses comparativos, do *self-interest* e cinismo atribucional e do viés de negatividade, tal parece dever-se ao facto de possuímos uma propensão: para nos avaliarmos mais positivamente em comparação com os outros (e, em consequência, para avaliarmos os outros mais negativamente em comparação connosco); para atribuir motivações egoístas ao comportamento dos outros (e mesmo quando nos deparamos com comportamentos aparentemente bons, tendemos a reconstruí-los de forma negativa); e para conferir mais peso a informação com valência negativa do que a informação com valência positiva. Recorrendo a um paradigma de atualização de crenças, o presente trabalho vem complementar o conhecimento em torno do fenómeno, ao demonstrar que a manutenção deste cinismo em relação ao comportamento dos outros no domínio moral deve ser, pelo menos em parte, possibilitada por um processo de atualização de crenças assimétrica pessimista. Este processo é caracterizado por uma maior facilidade de assimilação de nova informação negativa ou pessimista acerca do comportamento dos outros do que de nova informação positiva ou otimista e opera particularmente quando as pessoas se deparam com comportamentos imorais, por oposição a morais. Este trabalho replica ainda evidências da existência de um efeito *better-than-average* (BTA) no domínio moral, mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais. A ausência de correlações significativas entre este efeito BTA e o efeito de atualização cínica de crenças não apoia uma abordagem motivacional ao fenómeno. São discutidas potenciais implicações destes resultados para a Psicologia Social e para a sociedade em geral, apontadas limitações do estudo e sugeridas possíveis direções futuras.

**Palavras-chave:** atualização de crenças; comportamento moral; cinismo; assimetria; *better-than-average*

## **Abstract**

Even though we largely live surrounded by people we presumably consider good from a moral standpoint, we tend to think negatively about humanity's moral behavior. According to the literature in Social Psychology regarding comparative biases, self-interest and attributional cynicism and the negativity bias, this appears to be due to the fact that we possess a general propensity: to evaluate ourselves more positively in comparison to others (and, as a result, to evaluate others more negatively in comparison to ourselves); to attribute selfish motives to other people's behavior (and even when we come across apparently good behavior, we tend to reconstrue it in a negative light); and to give more weight to negatively valenced information than positively valenced information. Using a belief updating paradigm, the present work complements knowledge around this phenomenon, demonstrating that the persistence of this cynicism concerning other people's behavior in the moral domain should be, at least in part, made possible by a process of pessimistic asymmetric belief updating. This process is defined by an easier assimilation of new negative or pessimistic information regarding other people's behavior than of new positive or optimistic information and operates particularly when people face immoral, as opposed to moral behavior. This work also replicates evidence of the existence of a better-than-average (BTA) effect in the moral domain, more pronounced in the case of immoral than moral behavior. The absence of significant correlations between this BTA effect and the cynical belief updating effect does not give support to a motivational approach to the phenomenon. Potential implications of these results to Social Psychology and society in general are further discussed, limitations of the study pointed out and possible future directions proposed.

**Keywords:** belief updating; moral behavior; cynicism; asymmetry; better-than-average

## Índice

Introdução .....	1
Vieses Comparativos.....	2
<i>Self-Interest</i> e Cinismo Atribucional.....	9
Viés de Negatividade .....	15
O Presente Estudo .....	22
Problema de Investigação e Relevância .....	22
Objetivos, Paradigma Experimental e Hipóteses .....	24
Método .....	33
Participantes .....	33
Material .....	33
Procedimento.....	34
<i>Design</i> experimental.....	36
Resultados .....	37
Tarefa de Atualização de Crenças .....	37
Tarefa BTA .....	40
Relação Entre a Tarefa de Atualização de Crenças e a Tarefa BTA.....	40
Discussão .....	43
Implicações.....	48
Limitações e Direções Futuras .....	50
Conclusão.....	58
Referências.....	59
Anexo A – Lista Final de Comportamentos Seleccionados .....	72
Anexo B – Consentimento Informado e Instruções .....	75

## Introdução

Na perspectiva do senso comum, não deverá ser difícil recordar momentos do dia-a-dia em que nos deparámos com maus comportamentos alheios que nos fizeram sentir irados e desiludidos com a condição humana e que conduziram a considerações tais como “as pessoas só se preocupam com elas mesmas” ou “as pessoas são todas iguais” – por exemplo, um condutor que se atravessou numa fila de trânsito, um aglomerado de pessoas que dificultaram a nossa saída do metro ou alguém que se colocou à nossa frente numa fila de supermercado. Pelo contrário, parece muito menos provável que reflitamos acerca da bondade da humanidade quando testemunharmos boas ações, que tendem a ser consideradas atos isolados ou exceções. Curiosamente, uma simples pesquisa à data da escrita deste trabalho no motor de busca *Google* iniciada por “as pessoas são” revela de imediato “as pessoas são más” como uma das sugestões, sendo que estas sugestões são derivadas, em parte, das pesquisas mais populares e relevantes levadas a cabo pelos internautas.

Também a literatura tem apontado para esta tendência geral para pensar negativamente sobre o comportamento moral das outras pessoas. Brandimarte, Vosgerau, e Acquisti (2018), por exemplo, constatarem que as pessoas parecem atuar como juízes críticos do comportamento passado dos outros. Estes autores mostraram que enquanto os comportamentos positivos/morais recentes dos outros são muito mais influentes nas impressões que as pessoas formam acerca deles do que os comportamentos positivos/morais que ocorreram há muito tempo, a informação acerca dos seus comportamentos negativos/imorais não perde peso, independentemente de terem acontecido num passado próximo ou distante, o que promove julgamentos globalmente desfavoráveis das outras pessoas.

De forma aparentemente contraditória, outros estudos demonstraram que as pessoas acreditam que o “*self* verdadeiro” (o seu e o dos outros) é fundamentalmente bom do ponto de vista moral. Consideram, por exemplo, que o “*self* verdadeiro” se encontra mais refletido nas boas ações do que nas más ações das pessoas, que conhecem mais acerca do “*self* verdadeiro” de indivíduos altamente morais do que de outros indivíduos e, reciprocamente, que os indivíduos possuem mais traços morais quando sentem que conhecem o seu “*self* verdadeiro”, e que, no geral, existe dentro de cada indivíduo algo que o motiva a comportar-se de formas moralmente virtuosas (e.g., Christy, Kim, Vess, Schlegel, & Hicks, 2017; Newman, Bloom, & Knobe, 2014; Newman, De Freitas, & Knobe, 2015). No entanto, a maioria destes trabalhos examinou as crenças dos participantes acerca do “*self* verdadeiro” de alvos hipotéticos. Num estudo que comparou as crenças acerca do “*self* verdadeiro” do próprio e de outros reais



(Bench, Schlegel, Davis, & Vess, 2015), concluindo que os dois são vistos como comparavelmente favoráveis, os outros em questão diziam respeito a amigos próximos, o que pode explicar estas semelhanças. Além disso, mesmo que as pessoas considerem que o “*self* verdadeiro” dos outros se encontra mais refletido nas suas boas ações do que nas suas más ações, tal não implica necessariamente que acreditam que o comportamento habitual dos outros é mais moral (revelador do seu “*self* verdadeiro”) do que imoral. O contrário parece ser o caso.

Levanta-se, assim, uma questão intrigante: vivemos, em grande parte, rodeados de pessoas que presumivelmente consideramos boas do ponto de vista moral, nomeadamente as nossas famílias e grupos de amigos; no entanto, tendemos a pensar negativamente sobre o comportamento das outras pessoas em geral. Quais serão, então, os mecanismos subjacentes a este fenómeno? Porque é que se verifica uma tendência para o cinismo em relação ao comportamento moral da humanidade?

De acordo com investigação anterior no âmbito da Psicologia Social, existem 3 razões principais que podem estar subjacentes a este cinismo. A primeira diz respeito à literatura sobre vieses comparativos, segundo a qual as pessoas se consideram superiores aos outros, ou consideram que os outros são inferiores a si, numa variedade de domínios. A segunda refere-se à literatura sobre *self-interest* e cinismo atribucional. A última está relacionada com o conhecido viés de negatividade, que guia a atenção das pessoas para informação com valência negativa. Cada um destes tópicos será apresentado de seguida de forma mais aprofundada.

## **Vieses Comparativos**

Quando efetuam comparações, as pessoas apresentam uma tendência para se julgarem a si próprias de forma mais favorável do que os outros, um efeito demonstrado nomeadamente ao nível da avaliação de traços de personalidade e competências e denominado efeito *better-than-average* (BTA; Alicke, 1985; Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; para uma revisão, ver Alicke & Govorun, 2005) ou efeito *above-average* (Dunning, Meyerowitz, & Holzberg, 1989). De modo particular, é típico as pessoas reportarem possuir características positivas, como inteligência e educação, em maior grau do que os outros e características negativas, como desobediência e preguiça, em menor grau do que os outros (Alicke, 1985; Alicke et al., 1995; Brown, 1986; Dunning et al., 1989; Kruger & Gilovich, 2004).

Mas as diferenças eu-outros evidenciadas na literatura vão além da avaliação de traços de personalidade e competências. As pessoas também pensam que são melhores e, da mesma

forma, que os outros são piores numa variedade de outros domínios, desde a probabilidade de enviesamento até à capacidade de livre arbítrio, passando pela suscetibilidade a táticas persuasivas e pelo conformismo.

Especificamente, as pessoas acreditam que os outros são mais enviesados do que elas próprias, sendo mais vulneráveis a um conjunto de vieses cognitivos e motivacionais, como a alocação egoísta de responsabilidade por esforços coletivos, o efeito *halo* positivo ou o erro fundamental de atribuição (Kruger & Gilovich, 1999; Pronin & Kugler, 2007; Pronin, Lin, & Ross, 2002; para uma revisão, ver Pronin, Gilovich, & Ross, 2004) – um fenómeno que Pronin e colaboradores denominaram *bias blind spot*. Do mesmo modo, consideram, em geral, que os outros são menos objetivos, menos imparciais e menos perspicazes do que elas (Armor, 1999).

Acreditam também que os outros são mais suscetíveis a táticas persuasivas, detetando uma maior influência das comunicações de massa, como campanhas políticas ou representações de violência nos média, nas atitudes e comportamentos dos outros do que nas próprias atitudes e comportamentos (Davison, 1983; Douglas & Sutton, 2004; Innes & Zeitz, 1988).

Os outros são igualmente vistos como mais conformistas do que o próprio, sendo mais influenciados, por exemplo, pelas ações de outros espectadores na intervenção em emergências, por outros condutores na velocidade de condução ou pelos pares no envolvimento em comportamentos de risco (Pronin, Berger, & Molouki, 2007).

As pessoas apresentam também uma tendência para acreditar que os outros são menos livres do que elas, dado que as suas vidas são supostamente menos guiadas pelos princípios do livre arbítrio – os seus passados e futuros são mais previsíveis *a priori*, incluindo eventos como o término de uma relação passada ou o percurso de uma carreira futura; existem menos percursos possíveis nos seus futuros, incluindo possibilidades relacionadas com o local de residência, emprego e estilo de vida futuros; e o poder causal dos seus desejos e intenções pessoais sobre o seu comportamento é menor (Pronin & Kugler, 2010). Esta crença geral de que os outros são menos livres do que o próprio manifesta-se também através de uma tendência para acreditar que o comportamento dos outros é mais afetado pelos constrangimentos revelados no seu comportamento passado (Helzer & Dunning, 2012).

Relacionado com estas ideias, surge, como resultado do processo inferencial que as pessoas necessariamente conduzem quando se referem à mente dos outros, um viés segundo o qual tendem a acreditar que são mais capazes mentalmente do que os outros (Epley & Waytz, 2010; Schroeder, Waytz, & Epley, 2017; para uma revisão, ver Waytz, Schroeder, & Epley, 2014). Em particular, as mentes dos outros são vistas como menos sofisticadas, menos

racionais, menos complexas, menos capazes de pensamento deliberado e vontade consciente, menos profundas, menos intensas, com menos impacto causal ao nível do comportamento e menos objetivas do que a do próprio (Haslam, Bain, Douge, Lee, & Bastian, 2005; Waytz et al., 2014). As pessoas consideram mesmo que são mais essencialmente humanas do que os outros, possuindo características associadas a conceitos como a flexibilidade cognitiva e a responsividade emocional em maior grau do que as outras pessoas (Haslam et al., 2005).

Entre outros aspetos, as pessoas consideram ainda que os outros são piores condutores do que elas próprias, que são mais habilidosas e menos arriscadas no seu comportamento de condução (Svenson, 1981), independentemente do seu historial de acidentes automóveis sérios (Preston & Harris, 1965), que os outros são menos responsáveis por produtos conjuntos, por exemplo, fazer o pequeno-almoço num contexto de casal, sobrestimando as suas próprias contribuições (Ross & Sicoly, 1979), que a saúde dos outros é pior do que a sua, sendo menos saudáveis, ficando doentes mais vezes por ano e possuindo maior probabilidade de vir a contrair determinadas doenças (Larwood, 1978), que as relações românticas dos outros vão ter menor longevidade do que as do próprio (Epley & Dunning, 2006; MacDonald & Ross, 1999) e até, no caso específico de estudantes de gestão e de presidentes de empresas, que os outros possuem menos competências e são menos bem-sucedidos ao nível da gestão de vendas do que o próprio (Larwood & Whittaker, 1977).

Quando se trata de comparações de risco referentes a eventos de vida futuros, esta tendência para julgar o próprio de forma mais favorável do que os outros adquire o nome de otimismo irrealista (Weinstein, 1980), viés otimista (Weinstein, 1989) ou ilusão de invulnerabilidade exclusiva (*illusion of unique invulnerability*; e.g., Perloff & Fetzer, 1986), incluindo-se também na categoria de ilusões positivas (e.g., Taylor & Brown, 1988). Este efeito robusto e pervasivo reflete-se numa subestimação da probabilidade de o próprio vir a experienciar eventos de vida negativos, como divorciar-se, sofrer um ataque cardíaco ou ser despedido, em comparação com os outros (Perloff & Fetzer, 1986; Weinstein, 1980), e numa sobrestimação da probabilidade de o próprio vir a experienciar eventos de vida positivos, como receber uma boa proposta de emprego ou viver além dos 80 anos, em comparação com os outros (Weinstein, 1980).

De modo mais relevante para o presente propósito, também ao nível dos julgamentos comparativos no domínio moral se verifica um efeito semelhante ao efeito BTA geral documentado em domínios não morais, como aqueles até aqui referidos. Isto é, as pessoas acreditam que são melhores do ponto de vista moral do que os outros e, simultaneamente, que os outros são piores do ponto de vista moral do que elas próprias. Esta predisposição no

domínio moral era já evidente no estudo clássico de Alicke (1985) com traços de personalidade, no qual os participantes se consideraram, por exemplo, mais cooperativos, atenciosos, amáveis, leais e sinceros do que a pessoa típica, mas menos agressivos, traiçoeiros, indelicados, maus e antiéticos. Também Kruger e Gilovich (2004) mostraram que as pessoas se consideram, entre outros atributos, mais justas, generosas e morais do que os outros.

Mas, para além dos traços de personalidade, este efeito BTA no domínio moral também se verifica de forma bastante robusta no âmbito comportamental. As pessoas consideram que os outros tendem a agir de forma menos pró-social do que elas mesmas, comportando-se menos generosa, altruísta, solidária, cooperativa, ética e bondosamente. Na realidade, pode dizer-se que o efeito BTA é especialmente forte no que diz respeito às questões do comportamento moral e ético (Fetchenhauer & Dunning, 2006), na medida em que, por exemplo, as pessoas parecem estar dispostas a admitir em determinadas circunstâncias que os outros são mais inteligentes do que elas, mas são altamente relutantes a aceitar que os outros possam ser mais pró-sociais do que elas (Allison, Messick, & Goethals, 1989; van Lange, 1991; van Lange & Sedikides, 1998). Especificamente, Allison e colaboradores (1989) e van Lange (1991) mostraram que as pessoas acreditam possuir maior probabilidade de desempenhar comportamentos desejáveis e menor probabilidade de desempenhar comportamentos indesejáveis do que os outros, mas que a magnitude deste efeito é maior no domínio da moralidade do que no domínio da inteligência. Isto é, a assimetria eu-outros é mais forte no caso de comportamentos bons/maus (morais/imorais) do que no caso de comportamentos inteligentes/pouco inteligentes, um fenómeno que os autores denominaram efeito Muhammad Ali. Já van Lange & Sedikides (1998) apresentaram evidências da generalização deste efeito com participantes holandeses, comprovando a sua estabilidade intercultural.

De facto, segundo Klein e Epley (2016, 2017), existem poucos vieses ao nível do julgamento humano que sejam mais fáceis de demonstrar do que a tendência para acreditar que o próprio é mais moral do que os outros. Os autores apresentaram diversas evidências desta crença no nível comportamental. Nomeadamente, os participantes dos seus estudos acreditaram que os outros possuíam maior probabilidade do que o próprio de bater num carro estacionado e fugir sem deixar um bilhete, de roubar uma gorjeta de 20 dólares destinada a um empregado, de se envolver num caso extraconjugal, de mentir aos colegas de trabalho para obter uma promoção, de vender um produto a um preço inflacionado, de se apressar para ficar com um lugar de autocarro em vez de uma senhora idosa e de oferecer ajuda desonestamente. Ademais, os participantes recordaram mais comportamentos positivos para si próprios do que para os outros e as ações imorais dos próprios foram avaliadas como menos negativas do que as ações

imorais dos outros – isto é, as pessoas lembraram-se de os outros serem mais imorais no passado do que elas próprias teriam sido (Klein & Epley, 2016). Os participantes de Klein e Epley (2017) também apresentaram maior probabilidade de fazer inferências de caráter negativas a partir do comportamento não ético dos outros do que a partir do seu próprio comportamento não ético, e acreditaram ainda que os outros não se sentem tão mal quanto o próprio no rescaldo de uma ação não ética e que os outros são mais capazes de desempenhar comportamentos não éticos extremos do que o próprio.

Já Epley e Dunning (2000), num estudo no qual foi pedido para preverem o próprio comportamento e o comportamento dos seus pares em situações com implicações morais ou altruístas, verificaram que os participantes acreditaram que os seus pares possuíam menor probabilidade do que eles próprios de comprar flores para ajudar uma associação de caridade, de cooperar num jogo com o “dilema do prisioneiro”, de doar parte do dinheiro recebido pela participação num estudo experimental a uma organização sem fins lucrativos e de se voluntariar para uma tarefa experimental longa e desagradável no lugar de um parceiro.

Os participantes de um trabalho de Goethals (1986) também consideraram que os seus pares possuíam menor probabilidade do que os próprios de desempenhar uma variedade de comportamentos morais cooperativos, como devolver 5 dólares extra que receberam num troco ou parar de estudar para um exame final para transportar um amigo ao aeroporto, e maior probabilidade do que os próprios de desempenhar comportamentos imorais egoístas, como comer a maior das duas últimas fatias de *pizza*. Do mesmo modo, consideraram que os seus pares possuíam menor probabilidade do que o próprio de se comportar cooperativamente num jogo com o “dilema do prisioneiro” e de doar sangue num evento de caridade.

Noutros estudos, Balcetis e Dunning (2008) constataram que os participantes sobrestimaram a respetiva generosidade em comparação com a generosidade dos seus pares, prevendo que os pares iriam comprar menos bolos numa venda solidária para ajudar um hospital na Tanzânia, comprar menos lírios numa venda solidária para ajudar os Jogos Paralímpicos, comprar menos flores numa venda solidária para angariar dinheiro para a *American Cancer Society* e apoiar menos um evento de caridade que promovia a investigação em torno do cancro. Num trabalho posterior dos mesmos autores, os participantes estimaram que os outros iriam ajudar menos um experimentador a apanhar peças de *puzzle* que caíram no chão do que o próprio e ainda que os outros iriam doar menos do dinheiro que receberam como crédito num estudo experimental a uma instituição de caridade na Tanzânia do que o próprio (Balcetis & Dunning, 2013). No geral, estes participantes previram que eles próprios iriam

escolher a opção pró-social cerca de 90% do tempo, mas que, em média, apenas 47% dos seus pares se iriam envolver nos mesmos comportamentos pró-sociais.

As pessoas consideram ainda que os outros possuem menor probabilidade do que elas de oferecer o seu lugar num autocarro lotado a uma mulher grávida (Goethals, Messick, & Allison, 1991) e de exercer o comportamento de voto (Epley & Dunning, 2006).

Um outro domínio moral no qual o efeito BTA se torna evidente é o domínio da justiça. Messick, Bloom, Boldizar, e Samuelson (1985) pediram aos seus participantes que escrevessem listas de comportamentos justos e injustos levados a cabo pelos próprios ou pelos outros, e encontraram uma tendência pervasiva dos participantes para se associarem a si próprios com comportamentos justos e aos outros com comportamentos injustos. Os autores perceberam, em particular, que as pessoas acreditam que são mais justas do que os outros porque pensam que os outros se envolvem em comportamentos justos menos frequentemente do que elas e em comportamentos injustos mais frequentemente do que elas, ainda que, de modo geral, considerem que tanto o próprio como os outros levam a cabo mais frequentemente comportamentos justos do que injustos. Verificaram também que as subcategorias de comportamentos injustos que os participantes associaram com os outros foram diferentes das associadas com o próprio – os comportamentos injustos que os participantes associaram consigo disseram quase exclusivamente respeito a atos de desconsideração interpessoal, enquanto os comportamentos injustos que associaram com os outros incluíam atos de desconsideração interpessoal, mas também atos ilegais e comportamentos caracterizados por ganância económica. Liebrand, Messick, e Wolters (1986) apresentaram evidências da estabilidade ou generalização cultural deste viés de justiça com participantes falantes de holandês nos Países Baixos. Estes autores verificaram ainda que os comportamentos injustos associados aos outros foram considerados particularmente salientes e foram recordados mais frequentemente do que os restantes.

Diversos estudos têm também demonstrado que as pessoas consideram que os outros são mais preconceituosos do que elas próprias, sendo que mesmo quando a maioria das pessoas não apoia explicitamente opiniões preconceituosas, elas acreditam frequentemente que a maioria dos outros apoia essas opiniões (e.g., Judd, Park, Yzerbyt, Gordijn, & Muller, 2005; O'Gorman, 1975; O'Gorman & Garry, 1976). Judd e colaboradores (2005), por exemplo, verificaram que as pessoas acreditam que os outros possuem crenças intergrupais (viés avaliativo intergrupar e homogeneidade do *outgroup*) mais enviesadas do que as que elas próprias adotam, tanto em contextos intergrupais baseados na etnia, como em contextos intergrupais baseados no género e na nacionalidade, e tanto quando as próprias pessoas exibem

estas crenças intergrupais enviesadas (contexto de nacionalidade), como quando as próprias pessoas não exibem crenças intergrupais enviesadas (contextos de etnia e género).

De acordo com Heath (1999), as pessoas acreditam ainda que os outros são mais motivados por incentivos situacionais ou extrínsecos (aqueles que envolvem aspetos de uma situação que seriam facilmente verificados por juízes independentes; por exemplo: segurança no trabalho, salário, benefícios, elogios do diretor) do que elas próprias, e menos motivados por incentivos disposicionais ou intrínsecos (aqueles que envolvem estados internos difíceis de verificar por juízes independentes; por exemplo: uma mudança interna no conhecimento, como “aprender coisas novas” ou desenvolver capacidades, ou um estado interno de satisfação, como “sentir-se bem consigo mesmo”) do que elas próprias.

Por último, também a literatura relacionada com julgamentos acerca do comportamento passado tem evidenciado importantes diferenças eu-outros no domínio moral. Steimer, Mata e Simão (2019) verificaram que as pessoas tendem a considerar as suas memórias pessoais de comportamentos passados positivos como mais significativas e mais reveladoras da pessoa que são atualmente do que as memórias de comportamentos passados negativos, fazendo inferências acerca da sua personalidade atual a partir de comportamentos passados positivos, mas não a partir de comportamentos passados negativos, sendo que este efeito da valência não se verifica quando se trata de comportamentos passados dos outros, sejam eles conhecidos ou estranhos. No caso dos outros, o comportamento passado é tido como diagnóstico da personalidade presente, sendo possível fazer inferências acerca dos traços presentes a partir desse comportamento, seja ele positivo ou negativo. Da mesma forma, o efeito da valência sobre o significado percebido foi mediado pela desejabilidade dos comportamentos, sendo que as pessoas atribuíram tanto mais peso aos seus comportamentos passados quanto mais desejáveis esses comportamentos eram, mas apenas para o próprio, não para os outros. De modo semelhante, Ross & Wilson (2002) tinham já apresentado evidências de um viés de distanciamento subjetivo segundo o qual as pessoas reportam sentir-se mais distantes de “eus” e experiências antigas com implicações desfavoráveis para a sua visão atual de si próprias do que de “eus” e experiências igualmente distantes com implicações lisonjeiras, mas apenas no caso de eventos pessoais passados e não no caso de eventos passados da vida de outros conhecidos.

Na verdade, o facto de as pessoas acreditarem que são melhores do que os outros não significa necessariamente que considerem que os outros são maus – os outros podem até ser bons, mas o próprio é tido como ainda melhor (ou ainda menos mau). Efetivamente, a forma como as pessoas avaliam os outros quando os comparam consigo pode informar mais sobre o

modo como se encaram a si próprias do que sobre o modo como encaram os outros. De acordo com Douglas e Sutton (2004), é possível que não exista, na realidade, um *locus* de erro invariante no que diz respeito aos vieses eu-outros, sendo que por vezes estes vieses devem emergir a partir de erros acerca do próprio e outra vez a partir de erros acerca dos outros. Os autores aconselham, então, que se analise a literatura referente ao domínio em questão.

No domínio do comportamento moral, em particular, parece que os vieses comparativos são mais frequentemente decorrentes de erros em relação ao próprio do que de erros em relação aos outros, sendo este o caso em muitas das áreas até aqui referidas, nas quais se verifica uma sobrestimação da probabilidade de o próprio desempenhar comportamentos moralmente desejáveis, mas previsões relativamente precisas em relação aos outros (Balcetis & Dunning, 2008, 2013; Epley & Dunning, 2000, 2006), ou uma subestimação da probabilidade de o próprio desempenhar comportamentos moralmente indesejáveis, mas previsões mais realistas em relação aos outros (Klein & Epley, 2016, 2017). Segundo Epley e Dunning (2006), esta dificuldade em prever com precisão o próprio comportamento, mas prever com relativa precisão o comportamento dos outros é, de facto, mais bem ilustrada no âmbito dos comportamentos socialmente ou moralmente desejáveis.

Ainda assim, alguns autores apresentam evidências de subestimação da moralidade dos outros ou de sobrestimação da imoralidade dos outros (Diekmann, Samuels, Ross, & Bazerman, 1997; Fetchenhauer & Dunning, 2009; Miller, 1999; Miller & Ratner, 1998). E independentemente de o viés comparativo ter origem em erros relativamente ao próprio ou em erros relativamente aos outros, o facto de as pessoas se considerarem melhores (ou menos más) do que os outros do ponto de vista moral contribui para uma perceção negativa dos outros, pelo menos em comparação com o próprio. Ou seja, em termos relativos, os outros parecem, de facto, moralmente maus.

### ***Self-Interest* e Cinismo Atribucional**

Segundo a literatura, parece que as pessoas apresentam geralmente uma tendência para sobrestimar o poder do *self-interest*, nomeadamente ao nível das atitudes, dos julgamentos e do comportamento dos outros. De modo particular, Miller (1999) propôs a existência de uma norma do *self-interest* nas culturas ocidentais, habitualmente mais individualistas, que especifica que o *self-interest* é e deve ser um poderoso determinante do comportamento. Isto é, o autor defende que as pessoas são socializadas de forma a assumir que o *self-interest* é um guia razoável do comportamento, tanto a nível descritivo – as pessoas aprendem a assumir que



a maioria das outras pessoas é egoísta, agindo de acordo com o seu *self-interest* –, como a nível prescritivo – o próprio deve também reger-se pelos princípios do *self-interest*, sendo esse o comportamento socialmente desejável. Desta forma, o *self-interest* é concebido como uma norma cultural e, assim, pode dizer-se que as pessoas tendem a subestimar a moralidade dos outros.

Esta ideia foi, em parte, baseada em investigação empírica que demonstrou que as pessoas sobrestimam, por exemplo, o impacto que uma recompensa financeira exerce na disponibilidade dos seus pares para desempenhar um ato pró-social como o de doar sangue (Miller & Ratner, 1998). Miller e Ratner (1998) constataram que também o impacto da pertença grupal nas atitudes dos outros em relação a políticas que afetam o respetivo grupo tende a ser sobrestimado. Por exemplo, os participantes de um dos seus estudos esperavam que as mulheres possuíssem atitudes mais favoráveis do que os homens relativamente à cobertura do aborto no plano de saúde do governo, apesar de as atitudes de homens e mulheres em relação a tal proposta não terem, na realidade, diferido. Para além do impacto do género nas atitudes em relação a políticas em torno do aborto, os autores salientaram ainda o impacto do ano escolar nas atitudes em relação a políticas relacionadas com o álcool no *campus* universitário, o impacto do estatuto de fumador ou não fumador nas atitudes em relação aos impostos sobre o tabaco e a restrições ao fumo, e o impacto do interesse pessoal na disponibilidade para oferecer apoio público, mais especificamente para expressar oposição a uma ação governamental pendente que alegadamente ameaçava ou não ameaçava o estado de saúde do respetivo género. Notavelmente, a tendência para sobrestimar o impacto do *self-interest* nos outros nestes estudos estava, em grande parte, não relacionada com o seu impacto nas atitudes e comportamentos dos próprios participantes. Ou seja, mesmo quando as atitudes das pessoas em relação a determinada política social são incongruentes com o seu nível reconhecido de interesse nessa política, elas continuam a considerar que as atitudes dos outros serão congruentes com o respetivo *self-interest* (Miller & Ratner, 1998).

Já a componente prescritiva da norma do *self-interest* encontra-se bastante evidente num trabalho posterior de Ratner e Miller (2001), no qual os autores verificaram que as pessoas se sentem inibidas de se envolver em ações sociais incongruentes com o seu *self-interest* aparente, por receio de parecerem comportar-se de forma considerada desviante. Mais especificamente, os participantes previram que seriam avaliados negativamente se agissem em nome de uma causa que não os afetasse diretamente (por exemplo, homens a favor de políticas relacionadas com o aborto) ou de uma causa que os afetasse diretamente, mas na qual possuíssem atitudes incongruentes com o seu suposto *self-interest* (por exemplo, mulheres

contra políticas relacionadas com o aborto). E esta posição foi legitimada, na medida em que outros participantes reportaram surpresa e raiva quando um alvo agiu dessa forma. Ademais, os indivíduos sentiram-se mais confortáveis ao envolver-se em ações sociais e esperaram que os outros respondessem mais favoravelmente às suas ações se o assunto em causa fosse descrito como mais relevante para o seu próprio sexo do que para o sexo oposto. E os autores verificaram ainda que proporcionar um direito psicológico de agir aos indivíduos não afetados diretamente por uma causa fez com que estes apresentassem a mesma probabilidade que os indivíduos diretamente afetados de se envolver em ações sociais. Parece, então, que a norma do *self-interest* é tão prevalente que a expectativa de que as atitudes e comportamentos das pessoas se conformem ao seu *self-interest* faz com que, quando tal não se verifica, as atitudes e comportamentos que violam a norma provoquem desconforto no ator e confusão e raiva nos observadores (Ratner & Miller, 2001).

Outros resultados congruentes com a norma do *self-interest* defendida por Miller e colaboradores (Miller, 1999; Miller & Ratner, 1998; Ratner & Miller, 2001) tinham já sido avançados por Fein (1996) e Diekmann e colaboradores (1997), e foram também apresentados posteriormente por autores como Epley, Caruso, e Bazerman (2006), Fetchenhauer e Dunning (2009), e Ham e Vonk (2011).

Fein (1996) mostrou que as pessoas suspeitam muitas vezes de possíveis segundas intenções ou motivações ocultas no comportamento dos outros, sendo que esta predisposição pode mesmo ajudá-las a evitar vieses robustos como o erro fundamental de atribuição. Segundo o autor, tal acontece porque esta suspeição, definida como um estado dinâmico no qual o indivíduo entretém ativamente múltiplas hipóteses, possivelmente rivais, acerca das motivações ou da genuinidade do comportamento de uma pessoa, que implica que o indivíduo mantenha a noção de que o alvo está a tentar esconder algo, desencadeia um estilo de pensamento atribucional relativamente elaborado, consciente e deliberativo, caracterizado por uma consideração ativa e cuidadosa das potenciais motivações e causas que podem influenciar o comportamento. Este trabalho evidenciou, assim, a sensibilidade e sofisticação das pessoas no que diz respeito à deteção de *self-interest* nos outros.

Ao contrário de Fein (1996), que apontou para uma suspeita de segundas intenções ou motivações ocultas que envolve uma análise atribucional sofisticada, com pensamento consciente e deliberativo, Ham e Vonk (2011) argumentaram que as inferências acerca de segundas intenções ou motivações ocultas podem ser ativadas de forma espontânea, tão espontaneamente quanto outras inferências de traços. Quando os participantes dos seus estudos leram pequenas descrições comportamentais nas quais o ator era apresentado como dependente

do alvo, denotando uma possível motivação para o impressionar ou agradar (por exemplo: “O João voluntariou-se para ajudar a pintar a casa do seu chefe”, que pode sugerir “insinuante” ou o traço correspondente “prestativo”), eles apresentaram efeitos de facilitação da memória para traços referentes à motivação oculta implícita. Em contraste, quando o mesmo comportamento foi dirigido a um alvo neutro, em relação ao qual o ator não era dependente, de tal forma que a possibilidade de segundas intenções ou motivações ocultas era menos saliente (por exemplo: “O João voluntariou-se para ajudar a pintar a casa de um amigo”, que pode sugerir o traço correspondente “prestativo”, mas não tanto o traço referente à motivação oculta “insinuante”), apenas foram obtidas evidências para a inferência espontânea do traço correspondente e não da motivação oculta. Os autores perceberam, então, que as inferências acerca de motivações ocultas podem ser formadas sem uma intenção explícita de o fazer, sem muito tempo para pensar, sem muito esforço cognitivo, sem qualquer controlo sobre o processo e sem se estar ciente disso, tratando-se de um tipo de inferência espontânea que pode, assim, ser ativado numa fase precoce e automática do processo de atribuição. Numa tentativa de conciliar o seu trabalho com o de Fein (1996), Ham e Vonk (2011) avançaram que, provavelmente, as inferências acerca de comportamentos com possíveis segundas intenções ou motivações ocultas envolvem: uma etapa automática, espontânea, como indicado pelos seus resultados, na qual os traços e as motivações relacionadas com as possíveis segundas intenções são ativados juntamente com as inferências correspondentes; e uma etapa controlada, sugerida por resultados como os de Fein (1996), na qual os dois tipos de inferências são ponderados um em relação ao outro. De modo mais relevante para o presente propósito, ambos os trabalhos sugerem uma propensão, mais deliberativa ou mais automática, para a deteção de segundas intenções ou motivações ocultas no comportamento dos outros. É ainda de salientar que a compreensão de segundas intenções, possivelmente reveladoras de *self-interest*, no âmbito de julgamentos avaliativos de atos pró-sociais parece desenvolver-se aos poucos tão cedo quanto entre os 6 e os 10 anos de idade (Heyman, Barner, Heumann, & Schenck, 2014).

Também Diekmann e colaboradores (1997), que se propuseram a explorar a tensão entre o *self-interest* e a norma da igualdade, tinham já revelado que as pessoas consideram que os outros são particularmente parciais, apresentando, nomeadamente, um maior viés de favoritismo do *ingroup* do que eles próprios, tanto ao nível da alocação de recursos, como ao nível da avaliação da justiça de alocações desiguais que favoreciam os outros. Mais uma vez, tal aponta para a referida tendência para sobrestimar o impacto do *self-interest* nos outros.

Mais tarde, Epley e colaboradores (2006) perceberam que, sobretudo em contextos competitivos, levar as pessoas a considerar os pensamentos e perspetivas dos outros membros

do seu grupo pode reduzir julgamentos egocêntricos de que o próprio merece mais do que a sua parte justa de recursos grupais, mas tende a aumentar o respetivo comportamento egoísta, dado que considerar as perspectivas dos outros ativa teorias egoístas acerca do seu comportamento provável, conduzindo as pessoas a contrariar isso ao comportarem-se elas próprias mais egoistamente. Este padrão foi encontrado numa variedade de interações sociais e com referência a recursos grupais muito distintos, desde peixe a subsídios em dinheiro, passando por pepitas de chocolate e lucro corporativo. Pode, então, dizer-se que a incongruência entre os julgamentos mais justos e o comportamento injusto neste contexto deve-se, precisamente, ao facto de as pessoas assumirem que o *self-interest* possui um forte impacto no comportamento dos outros.

Já Fethenauer e Dunning (2009) constataram, recorrendo ao paradigma do jogo de confiança, que as pessoas subestimam sistematicamente a confiabilidade dos outros. Foi pedido aos participantes que decidissem ceder ou não ceder uma quantia de dinheiro a um indivíduo anónimo que poderia devolver ainda mais dinheiro ou ficar com tudo. Os participantes confiaram pouco, subestimando grandemente a proporção dos seus pares que iria devolver dinheiro. Na realidade, uma grande maioria dos participantes, por vezes tão alta quanto 90%, decidiu honrar a confiança e devolver dinheiro a alguém que tinha investido em si; no entanto, esta proporção foi subestimada cerca de 30 a 35%, dependendo do estudo. Tal revela elevados níveis de cinismo e aponta, mais uma vez, para a ideia de que o comportamento dos outros é frequentemente considerado fruto do respetivo *self-interest*, mesmo que não seja essa a realidade.

Também alguns dos trabalhos já mencionados no âmbito da literatura sobre vieses comparativos, como os de Kruger e Gilovich (1999) e de Pronin e colaboradores (2002), sugerem que as pessoas tendem a sobrestimar o impacto do *self-interest* nos julgamentos levados a cabo pelos outros. Os participantes de Kruger e Gilovich (1999), em particular, esperaram que os outros fossem egoístas nos seus julgamentos, nomeadamente na alocação de responsabilidade por esforços coletivos, independentemente de terem realmente sido ou não, sendo que mesmo quando os outros foram egoístas, não o foram tanto quanto esperado. Os autores denominaram este fenómeno cinismo ingénuo. Já os participantes de Pronin e colaboradores (2002) apresentaram uma tendência para indicar que o pensamento dos outros é afetado por vieses motivacionais, incluindo vieses caracterizados por pensamento *self-interested*.

A crença de que o *self-interest* é um importante determinante do comportamento é forte ao ponto de persistir mesmo em face de evidências supostamente desconfirmatórias, tal como

defendido por Critcher e Dunning (2011). Segundo os autores, isto acontece porque quando as pessoas refletem acerca de comportamentos altruístas tendem a reconstruí-los em termos de motivações egoístas, sendo que, pelo contrário, os comportamentos egoístas não são escrutinados da mesma forma em termos de motivações altruístas. Este padrão foi designado de cinismo atribucional. De modo particular, parece que as pessoas reagem a nova informação de que o comportamento altruísta é comum, interpretando-o como mais indicativo de *self-interest*. Nomeadamente, os participantes reagiram à notícia de que a maioria das pessoas é honesta passando a considerar a honestidade mais como um produto de *self-interest*; ou seja, ao invés de reverem as suas crenças acerca da prevalência global de egoísmo no mundo, os participantes decidiram que comportamentos aparentemente altruístas deviam ser, na verdade, egoístas, isto é, reviram o comportamento que acreditavam que o *self-interest* motivava. Perceberam também que, de acordo com uma análise Bayesiana, o mero pensamento atribucional leva as pessoas a detetar demasiado *self-interest* em ações aparentemente altruístas, tendo em conta as suas crenças prévias, enquanto detetam de forma precisa a quantidade prevista de *self-interest* em ações aparentemente egoístas. Constaram, assim, que o cinismo ocorre mesmo sem a apresentação de nova informação acerca da prevalência de determinados comportamentos, acentuando a ideia da existência de um cinismo excessivo. Por último, os participantes de um dos estudos de Critcher e Dunning (2011) leram acerca de filantropos reais cujos atos de generosidade tinham sido apresentados por importantes meios de comunicação. À medida que os participantes passaram mais tempo a considerar a razão pela qual tal filantropia foi desempenhada, formaram impressões mais cínicas das motivações dos filantropos. De notar ainda que os participantes apelaram mais frequentemente a motivações egoístas do que altruístas quando explicaram as ações dos filantropos. Este resultado atesta a robustez do efeito. Critcher e Dunning (2011) concluíram que as pessoas mantêm uma crença no poder do *self-interest* mesmo quando observam evidências que deveriam desafiar essa crença porque, de acordo com um padrão de cinismo atribucional, respondem de forma assimétrica a ações aparentemente altruístas e egoístas. Mesmo quando confrontadas com ações aparentemente altruístas, as pessoas reinterpretam-nas e insistem em atribuir motivações egoístas, isto é, *self-interest*.

Kim (2014) veio introduzir alguma incerteza no que diz respeito ao *self-interest*, considerando que a investigação em torno deste conceito é complexa, ambígua e até inconsistente. A autora pensa, por um lado, que as definições de *self-interest* adotadas nas ciências sociais são muito diversas, contribuindo para a percepção de que todos os comportamentos acabam por estar de acordo com o conceito, e acredita que as circunstâncias

nas quais o *self-interest* afeta o comportamento devem ser reconsideradas, levando à discussão e adoção de uma definição comum. Por outro lado, defende que a investigação tem avançado diversos casos nos quais o comportamento humano é inesperado e não se conforma com os princípios do *self-interest*. Ainda assim, Kim (2014) admite sempre que o *self-interest* é amplamente aceite como uma das motivações mais poderosas ao nível do comportamento humano, tanto por parte dos académicos como por parte das pessoas leigas. E se as pessoas são realmente cínicas em relação às motivações dos outros, acreditando que o seu comportamento é fortemente guiado por *self-interest* (Diekmann et al., 1997; Epley et al., 2006; Fein, 1996; Fetchenhauer & Dunning, 2009; Ham & Vonk, 2011; Heyman et al., 2014; Kruger & Gilovich, 1999; Miller, 1999; Miller & Ratner, 1998; Pronin et al., 2002; Ratner & Miller, 2001), até mesmo quando se deparam com ações aparentemente altruístas (Cricher & Dunning, 2011), então faz sentido que assumam frequentemente o pior acerca dos outros. Tal deve contribuir para uma perceção negativa do comportamento moral das pessoas em geral, que, desta forma, não se verifica apenas em termos relativos, em comparação com o próprio, mas também em termos absolutos.

### **Viés de Negatividade**

A literatura sugere a existência de um efeito de negatividade ou viés de negatividade prevalente, segundo o qual a informação negativa, como eventos, objetos, comportamentos ou traços pessoais, apresenta globalmente maior poder em comparação com a informação positiva. De modo particular, Rozin e Royzman (2001) propuseram que este viés de negatividade pode manifestar-se de 4 formas diferentes: a) potência negativa (*negative potency*) – as entidades negativas são mais fortes do que as entidades positivas equivalentes, isto é, são mais potentes (subjettivamente) em relação à sua magnitude objetiva; b) gradientes negativos mais íngremes (*steeper negative gradients*) – a negatividade dos eventos negativos cresce mais rapidamente com a aproximação a esses eventos no espaço ou no tempo do que a positividade dos eventos positivos; c) dominância negativa (*negative dominance*) – combinações de entidades positivas e negativas produzem avaliações que são mais negativas do que o previsto pela soma algébrica das valências subjetivas individuais, sendo esta a subclasse de fenómenos de viés de negatividade que os autores consideram ser mais robusta e informativa, e conferindo, neste contexto, destaque para o maior contágio das entidades negativas em comparação com as entidades positivas; e d) diferenciação negativa (*negative differentiation*) – as entidades negativas são mais variadas, produzem representações conceptuais mais complexas e

envolvem um reportório de respostas mais amplo, isto é, são interpretadas como sendo mais elaboradas e mais diferenciadas.

Apesar de defenderem que o viés de negatividade não é universal, Rozin e Royzman (2001) apresentam, enquadrada nesta sua taxonomia, uma revisão de evidências para o efeito em vários domínios, nomeadamente ao nível: da excitação fisiológica; da sensação e da percepção; da atenção e saliência; da aprendizagem; da motivação; do humor; até certo ponto, da memória; do contágio; da tomada de decisão; do desenvolvimento; e da empatia.

Já Baumeister, Bratslavsky, Finkenauer, e Vohs (2001) atestaram também, com base numa revisão abrangente da investigação relevante, que o “mau” é mais forte do que o “bom”, de tal forma que eventos com valência negativa possuem um maior impacto ao nível do comportamento humano do que eventos com valência positiva do mesmo tipo, e que este princípio geral é adaptativo do ponto de vista evolutivo e é extremamente pervasivo, aplicando-se a uma diversidade de processos psicológicos.

Baumeister e colaboradores (2001) começaram por constatar que, na vida quotidiana, os eventos negativos possuem consequências mais fortes e mais duradouras do que eventos positivos comparáveis, sendo que as pessoas reagem mais fortemente a eventos negativos do que a eventos positivos. As evidências apresentadas pelos autores abrangem pequenos eventos quotidianos, exposição experimental breve e odores aversivos, mas também grandes eventos de vida e traumas. Os eventos negativos produzem mais emoção, apresentam maiores efeitos em medidas de ajustamento e possuem efeitos mais prolongados. Além disso, segundo eles, a inexistência de um equivalente positivo do conceito de trauma é, em si própria, um sinal de que eventos negativos únicos possuem frequentemente efeitos que são muito mais duradouros e importantes do que quaisquer resultados de eventos positivos únicos.

No que diz respeito ao contexto relacional, Baumeister e colaboradores (2001) indicaram que as relações próximas são afetadas de forma mais profunda e conclusiva por ações destrutivas do que por ações construtivas, por comunicações negativas do que por comunicações positivas, e pelo conflito do que pela harmonia, sendo que estes efeitos se estendem à satisfação conjugal e até à sobrevivência da relação. Em particular, a reciprocidade de respostas negativas parece constituir um poderoso determinante da deterioração e eventual término de relações próximas. Assim, os efeitos mais fortes dos eventos negativos em comparação com os positivos nas relações próximas podem ser observados nas interações do dia-a-dia, mas também em padrões mais gerais, no plano emocional, na resolução de problemas e ao nível da comunicação conjugal. E mesmo no âmbito de relações não tão próximas e de outras formas de contacto interpessoal, interações hostis ou conflituosas são consideradas mais

fortes e possuem efeitos maiores do que interações amigáveis e harmoniosas. De modo semelhante, a investigação sobre suporte social indica que comportamentos negativos e conflituosos na rede social de um indivíduo possuem efeitos mais fortes do que comportamentos positivos e solidários. De modo particular, parece que os aspetos cooperativos da rede social de um indivíduo possuem pouca ou nenhuma relação com a depressão, o bem-estar e a satisfação com o suporte social, enquanto os aspetos perturbadores e não cooperativos se relacionam com esses conceitos.

Baumeister e colaboradores (2001) avançam também que as emoções negativas são mais importantes e possuem efeitos mais fortes ao nível do processamento cognitivo e do comportamento do que as emoções positivas, e que os esforços de regulação emocional consistem geralmente numa tentativa de evitar emoções negativas, por oposição a induzir ou prolongar emoções positivas. Os autores acreditam, deste modo, que o desejo das pessoas de se libertarem de um estado de humor negativo é mais forte do que o desejo de se envolverem num estado de humor positivo. A preponderância de palavras para reportar emoções negativas, em contraste com a maior frequência das emoções positivas, também sugere que as emoções negativas são mais poderosas. Ademais, as pessoas tendem a recordar melhor as emoções negativas.

De acordo com Baumeister e colaboradores (2001), alguns padrões de aprendizagem sugerem igualmente que os correlatos de eventos negativos são aprendidos de forma mais rápida e eficaz do que os correlatos de eventos positivos correspondentes. Especificamente, a aprendizagem e o condicionamento são mais fortemente afetados por aspetos negativos do que positivos, mesmo quando a magnitude objetiva do positivo e do negativo é igualada de forma precisa.

Parece que até mesmo ao nível dos estilos parentais se deteta um viés de negatividade. Os autores salientaram que os efeitos negativos da parentalidade e de outras experiências ambientais nas crianças são preponderantes – a parentalidade negativa e os ambientes familiares negativos sobrepõem-se ao efeito das influências genéticas, por exemplo, no que diz respeito à inteligência, enquanto a parentalidade positiva e os ambientes familiares positivos não aparentam possuir o mesmo poder.

Em termos cognitivos, Baumeister e colaboradores (2001) destacam que a informação negativa recebe mais atenção e um processamento mais minucioso do que a informação positiva. Já no caso da memória, apesar de não parecer constituir a força dominante, a literatura fornece alguma evidência para o maior poder do negativo em comparação com o positivo. Sugere-se também que o cérebro responde mais fortemente a informação negativa do que a



informação positiva e que retém a memória de aspetos negativos mesmo quando a resposta comportamental já foi extinta.

De acordo com os autores, o viés de negatividade torna-se também evidente no que toca ao autoconceito, no sentido em que os indivíduos apresentam, em geral, maior motivação para evitar autodefinições negativas do que para reivindicar autodefinições positivas, ainda que as autoavaliações inflacionadas no sentido positivo, descritas na secção sobre vieses comparativos, possam dar a impressão contrária. De modo semelhante, os estereótipos negativos e as reputações negativas são adquiridos mais facilmente e são mais resistentes à desconfirmação do que os estereótipos positivos e as reputações positivas.

O *feedback* negativo, como *feedback* associado a insucesso ou rejeição, possui maior impacto e poder do que o *feedback* positivo, como *feedback* associado a sucesso ou aceitação. Segundo a revisão de Baumeister e colaboradores (2001), é possível que as respostas defensivas consigam por vezes minimizar o impacto a longo prazo do *feedback* negativo, mas quando estas defesas não se encontram a operar, o *feedback* negativo é mais forte do que o positivo. Além disso, acredita-se que os esforços maiores e mais pervasivos para minimizar o *feedback* negativo do que para maximizar o *feedback* positivo refletem uma assimetria motivacional que também reconhece que o *feedback* negativo é mais importante do que o positivo.

Por último, Baumeister e colaboradores (2001) acrescentam que até a má saúde possui um maior impacto na felicidade do que a boa saúde, e que a própria saúde é mais afetada pelo pessimismo, refletido na presença ou ausência de uma perspetiva negativa, do que pelo otimismo, refletido na presença ou ausência de uma perspetiva positiva. É ainda referido que o *stress* e a ausência de suporte social estão associados à imunodepressão, enquanto os seus opostos – o relaxamento e as melhorias ao nível do suporte social – não parecem possuir efeitos benéficos análogos.

Os trabalhos de Rozin e Royzman (2001) e Baumeister e colaboradores (2001) são essencialmente complementares, como assumido pelos próprios autores. A principal diferença diz respeito ao facto de os primeiros terem focado as interações entre a informação positiva e a informação negativa, enfatizando a tendência para os efeitos da informação negativa dominarem, por vezes por completo, os efeitos da informação positiva quando ambas se apresentam em conjunto numa única configuração, enquanto os últimos evidenciaram efeitos paralelos/independentes de fatores positivos e negativos no comportamento humano. Ambos concluíram que, de facto, a informação negativa é mais saliente, potente e dominante do que a informação positiva e produz, de modo geral, efeitos maiores, mais consistentes, mais eficazes,

mais multifacetados ou mais duradouros. E a pervasividade deste viés de negatividade encontra-se refletida no facto de que o mesmo é encontrado tanto ao nível da cognição como da motivação, tanto ao nível de processos internos e intrapsíquicos como de processos interpessoais, tanto em ligação com decisões acerca do futuro como, até certo ponto, com memórias do passado, e tanto na aprendizagem animal como no complexo processamento de informação humano e nas respostas emocionais.

De modo mais relevante para o presente propósito, ambas as revisões (Baumeister et al., 2001; Rozin & Royzman, 2001) apontaram para a existência bastante documentada e reconhecida pelos investigadores de um viés de negatividade robusto ao nível do processo de formação de impressões e do julgamento moral. Parece, então, que os comportamentos negativos ou imorais possuem mais peso e têm um impacto significativamente mais forte na impressão global que os indivíduos formam acerca dos outros do que os respetivos comportamentos positivos ou morais (e.g., Fiske, 1980; Skowronski & Carlston, 1987, 1992). Por exemplo, uma pessoa que rouba determinada quantia de dinheiro é considerada mais imoral do que uma pessoa que doa a mesma quantia de dinheiro é considerada moral.

No caso particular de Fiske (1980), a autora mostrou que os comportamentos negativos apresentam maior impacto do que a informação positiva na determinação da simpatia (*likability*) dos outros, sobretudo quando os comportamentos em causa são extremos. Uma medida de atenção (tempo de observação) replicou este padrão.

Já Skowronski e Carlston (1987, 1992) observaram um viés de negatividade no domínio moral, segundo o qual os comportamentos negativos do ponto de vista moral, nomeadamente comportamentos desonestos, apresentam maior influência no processo de formação de impressões do que os comportamentos positivos. Estes autores detetaram, porém, um viés de positividade no domínio da aptidão, no sentido em que neste contexto os comportamentos com maior impacto tendem a ser os comportamentos positivos, reveladores de inteligência e competência, ao invés dos negativos. Os autores enquadraram estes resultados numa teoria explicativa baseada na diagnosticidade dos comportamentos – no domínio moral, os comportamentos imorais são mais influentes porque são altamente informativos e, por isso, mais diagnósticos do que os comportamentos morais, dado que, à partida, apenas pessoas imorais se envolvem em comportamentos imorais, enquanto se espera que tanto pessoas morais como pessoas imorais se envolvam em comportamentos morais em determinadas circunstâncias; no domínio da aptidão, os comportamentos inteligentes ou competentes são mais influentes porque são altamente informativos e, por isso, mais diagnósticos do que os comportamentos pouco inteligentes ou pouco competentes, dado que, à partida, apenas pessoas

inteligentes ou competentes conseguem desempenhar comportamentos inteligentes ou competentes, enquanto se espera que tanto pessoas pouco inteligentes ou pouco competentes como pessoas inteligentes ou competentes desempenhem comportamentos pouco inteligentes ou pouco competentes em determinadas circunstâncias.

Martijn, Spears, Van der Pligt, e Jakobs (1992) reportaram resultados semelhantes: a informação comportamental negativa conduz a inferências mais seguras no domínio da moralidade, enquanto a informação comportamental positiva conduz a inferências mais seguras no domínio da aptidão. No entanto, ao combinarem traços do domínio da moralidade com traços do domínio da aptidão (moral positivo-aptidão positivo; moral positivo-aptidão negativo; moral negativo-aptidão positivo; moral negativo-aptidão negativo), estes autores encontraram fortes efeitos de negatividade em ambos os domínios, uma vez que todas as combinações de traços, à exceção da condição duplamente positiva (moral positivo-aptidão positivo), resultaram numa impressão mais negativa do que a simples média dos julgamentos referentes a traços de moralidade e a traços de aptidão isolados. Isto é, a informação negativa revelou-se mais influente na formação da impressão global do que a informação positiva, independentemente de dizer respeito ao domínio moral ou ao domínio da aptidão. Estes resultados replicaram o efeito de diagnosticidade avançado por Skowronski e Carlston (1987, 1992), mas mostraram que o maior poder da informação negativa em comparação com a informação positiva no âmbito da formação de impressões globais vai além da questão da diagnosticidade. Baumeister e colaboradores (2001) sugeriram, justamente, que apesar de terem sido propostas diversas teorias que podem explicar em parte a assimetria ao nível da formação de impressões, o maior poder da informação negativa parece persistir mesmo quando essas explicações não se aplicam.

Este viés de negatividade estende-se também ao processo de revisão de impressões iniciais. Desta forma, um comportamento imoral inconsistente com uma impressão inicial de outra pessoa tem o poder de alterar essa impressão e produz tempos de resposta relativamente longos, enquanto um comportamento moral inconsistente tem menos impacto na impressão (Reeder & Coover, 1986). Além disso, é necessária menos evidência comportamental para passar a considerar alguém uma má pessoa do ponto de vista moral do que para elevar alguém à categoria de boa pessoa (Klein & O'Brien, 2016). Enquanto os estudos anteriores enfatizaram a ponderação diferencial de diferentes tipos de informação que as pessoas aprendem acerca dos outros, isto é, que a informação negativa é ponderada mais fortemente do que a informação positiva, Klein e O'Brien (2016) revelaram um viés de negatividade ao nível da quantidade de

evidência necessária para alcançar julgamentos de caráter (o ponto no qual alvos neutros parecem tornar-se bons ou maus).

Na verdade, já Skowronski e Carlston (1992) tinham constatado que, no caso específico da dimensão moral, as implicações de comportamentos muito negativos são mais difíceis de contradizer com base em comportamentos positivos do que as implicações de comportamentos muito positivos com base em comportamentos negativos. Como notado pelos autores, de acordo com observação casual, pessoas que se envolveram em comportamentos extremamente imorais parecem ser frequentemente percebidas de forma negativa, mesmo que o seu restante comportamento seja exemplar. De facto, parece que enquanto as implicações de um ato incompetente podem ser compensadas por um único ato competente, as implicações de um ato imoral, ainda que possam ser eventualmente compensadas por benevolência suficiente e igualmente extrema, podem demorar uma vida a ser dissipadas (Skowronski & Carlston, 1992).

Também Risky e Birnbaum (1974) tinham anteriormente afirmado que a bondade geral de uma pessoa é maioritariamente determinada pela sua ação mais imoral, sendo que as boas ações possuem menos influência. Os autores admitiram que a adição de ações morais melhora as avaliações de conjuntos com itens imorais, mas esta compensação parece ser limitada. Assim, tendo cometido ações muito imorais, as pessoas não devem esperar conseguir redimir-se com boas ações.

Na sua revisão, Rozin e Royzman (2001) também contemplaram este viés de negatividade ao nível do julgamento moral. Os autores afirmaram que os atos imorais extremos possuem um efeito quase permanente, sendo frequentemente imperdoáveis, e mencionaram a noção de “*tragic flaw*” característica do drama Grego clássico, segundo a qual um único defeito numa personagem de outra forma admirável provoca a sua ruína. Mencionaram ainda o pequeno lapso ou a ação questionável única que arruinam uma carreira promissora. Por outro lado, apontaram que um crime como o homicídio não tende a ser equilibrado por um único ato heroico. E fizeram igualmente referência à fragilidade da pureza moral no sistema Hindu, que é difícil de atingir e manter, mas fácil de perder com base num só ato imoral.

Estes resultados complementam-se na comprovação da existência de um viés de negatividade geral ao nível da formação de impressões e do julgamento moral, desde as caracterizações iniciais que os indivíduos constroem acerca dos outros (e.g., Fiske, 1980; Martijn et al., 1992; Skowronski & Carlston, 1987, 1992) até à sua subsequente resistência (e.g., Klein & O’Brien, 2016; Reeder & Coover, 1986; Risky & Birnbaum, 1974; Skowronski & Carlston, 1992).

Autores como Lupfer, Weeks, e Dupuis (2000) defendem que o viés de negatividade no domínio moral, embora prevalente, não deve ser inevitável nem tão pervasivo quanto é frequentemente sugerido, indicando que o mesmo surge quando a informação acerca dos outros é inconsistente no que diz respeito às suas implicações morais, mas não quando a informação é consistente. Ainda assim, e tendo em conta as evidências apresentadas, pode dizer-se que, na generalidade, a informação negativa é mais influente do que a informação positiva num vasto conjunto de domínios que inclui o domínio moral, e que este viés deve, então, contribuir para uma perceção também ela negativa do comportamento moral das pessoas em geral.

## **O Presente Estudo**

### **Problema de Investigação e Relevância**

Parece, então, que tendemos a pensar negativamente sobre o comportamento das outras pessoas, apesar de vivermos, em grande parte, rodeados de pessoas que presumivelmente consideramos boas do ponto de vista moral, porque temos uma propensão geral para nos avaliarmos mais positivamente em comparação com os outros (e, em consequência, para avaliarmos os outros mais negativamente em comparação connosco), para atribuir motivações egoístas ao comportamento dos outros (e mesmo quando nos deparamos com comportamentos aparentemente bons, tendemos a reconstruí-los de forma negativa), e para conferir mais peso a informação com valência negativa do que a informação com valência positiva.

No entanto, é possível assumir que no dia-a-dia somos confrontados tanto com boas ações como com más ações por parte dos outros que nos rodeiam. E, na realidade, os outros podem até ser, de modo geral, maioritariamente bons. A título de exemplo, há que relembrar o trabalho de Fetscherin e Dunning (2009). Nos seus estudos, nos quais recorreram ao paradigma do jogo de confiança, uma grande maioria dos participantes, por vezes tão alta quanto 90%, decidiu honrar a confiança e devolver dinheiro a alguém que tinha investido em si. Deste modo, porque é que as pessoas não aprendem que os outros não são assim *tão* maus? Porque é que o cinismo em relação ao comportamento moral da humanidade se mantém?

Analisar os mecanismos subjacentes ao cinismo em relação ao comportamento moral da humanidade e, em particular, à manutenção destas crenças, aparentemente resistentes à aprendizagem, é relevante por diversas razões. Em primeiro lugar, focar a investigação ao nível do julgamento do comportamento no domínio moral é particularmente pertinente. Por um lado, porque qualquer tipo de previsão relacionada com o carácter e o comportamento dos outros no domínio moral detém uma clara importância para a vida social, na medida em que a bondade

do caráter e do comportamento de alguém determina o potencial dessa pessoa para ser prejudicial ou útil para o próprio (e.g., Fiske, Cuddy, & Glick, 2007; Goodwin, Piazza, & Rozin, 2014; Wojciszke, Bazinska, & Jaworski, 1998). Por outro lado, porque a moralidade é um conceito significativo para a questão da identidade pessoal (e.g., Goodwin et al., 2014). Ou seja, em teoria, e tal como sugerido por Goodwin e colaboradores (2014), a primazia do caráter e do comportamento moral deve refletir tanto considerações funcionalistas do ponto de vista social, como considerações mais simbólicas, baseadas na identidade. E, efetivamente, sabe-se que a moralidade é, na prática, um conceito dominante na percepção social, superando quase todas as outras características, incluindo a competência, quando os indivíduos formam impressões globais e avaliam os outros (e.g., Fiske et al., 2007; Goodwin et al., 2014; Wojciszke et al., 1998). O caráter moral é de tal forma importante que as pessoas preferem que os outros sejam morais, mesmo à custa de outras qualidades desejáveis (Goodwin et al., 2014). De facto, Wojciszke e colaboradores (1998) tinham já concluído que as categorias morais ocupam uma posição privilegiada nas impressões avaliativas globais que as pessoas formam acerca dos outros. Estes autores constataram que a informação relacionada com a moralidade é particularmente acessível do ponto de vista cognitivo, é mais procurada quando as pessoas reúnem informação para formar uma impressão global e influencia mais as impressões globais que as pessoas formam acerca dos outros, tanto quando os alvos são pessoas reais e conhecidas, como quando se trata de estranhos fictícios. E para além de ser o fator que mais importa aos indivíduos quando conduzem avaliações e formam impressões globais, a moralidade é também fundamental quando são realizadas avaliações específicas ou quando os indivíduos procuram conhecer os outros. Num trabalho de Hartley e colaboradores (2016), por exemplo, a moralidade foi o fator mais importante (em comparação com a competência e com a sociabilidade) para avaliações do quanto os participantes gostavam, respeitavam e conheciam alguém, tipos de julgamentos que são universais, cruciais para a interação social e levados a cabo com frequência no quotidiano.

Em segundo lugar, o estudo destes mecanismos deverá permitir uma compreensão mais aprofundada do fenómeno de cinismo em relação à bondade da humanidade e, consequentemente, informar possíveis custos deste cinismo, nomeadamente implicações ao nível do comportamento e das decisões sociais, no sentido em que qualquer processo que estimule o cinismo relativamente aos outros ou permita a sua manutenção pode também conduzir a uma redução das tendências pró-sociais e, eventualmente, contribuir para a intensificação do conflito humano. Se as pessoas considerarem que o comportamento dos outros é, em geral, moralmente mau e se não estiverem predispostas para aprender o contrário,

então não deverá ser inusitado que se verifique uma redução, por exemplo, da confiança, da cooperação e do compromisso com os pares, com consequências tanto nas decisões comuns do dia-a-dia, como em decisões menos frequentes e potencialmente mais importantes, como a de doar sangue ou a de contribuir com tempo ou dinheiro para uma causa social. De facto, tal como demonstrado nos já referidos estudos de Epley e colaboradores (2006), a ativação de teorias acerca do comportamento provavelmente egoísta dos outros tende a aumentar o comportamento egoísta do próprio, um efeito que os autores denominaram egoísmo reativo. Isto porque a expectativa de que os outros irão comportar-se de forma egoísta, aliada a um receio de poderem vir a ser exploradas por esse comportamento caso elas próprias atuem de forma pró-social, conduz as pessoas a contrariá-lo, comportando-se também elas mais egoistamente. Também Miller e colaboradores apontaram para a existência de uma espécie de profecia autoconfirmatória no contexto da norma do *self-interest* (Miller, 1999), segundo a qual a crença de que o *self-interest* afeta fortemente o comportamento dos outros incentiva as pessoas a comportarem-se igualmente de formas egoístas. Este tipo de profecia autoconfirmatória pode verificar-se para além do *self-interest*, ocorrendo de modo mais abrangente quando as pessoas consideram o comportamento dos outros no domínio moral. Em síntese, as crenças em relação ao comportamento dos outros podem influenciar a forma como o próprio decide comportar-se, e se as pessoas acreditam, em geral, que o comportamento dos outros é maioritariamente mau do ponto de vista moral, então é provável que também elas se comportem mais imoralmente em resposta.

### **Objetivos, Paradigma Experimental e Hipóteses**

Assim, o objetivo principal do presente estudo passa por tentar compreender de que forma as crenças negativas em relação ao comportamento moral da humanidade persistem. Para tal, propõe-se a existência de um 4º mecanismo potencialmente subjacente a este cinismo: a atualização assimétrica de crenças acerca do comportamento dos outros no domínio moral, ou atualização cínica de crenças.

Recentemente, Sharot, Korn, e Dolan (2011) propuseram um paradigma experimental desenvolvido com o propósito de estudar a atualização de crenças no âmbito da vasta literatura sobre otimismo irrealista. Como já referido, a tendência das pessoas para se julgarem a si próprias de forma mais favorável do que os outros estende-se também a comparações de risco referentes a eventos de vida futuros (e.g., Perloff & Fetzer, 1986; Taylor & Brown, 1988; Weinstein, 1980, 1989), um efeito que se reflete numa subestimação da probabilidade de o

próprio vir a experienciar eventos de vida negativos em comparação com os outros (Perloff & Fetzner, 1986; Weinstein, 1980) e numa sobrestimação da probabilidade de o próprio vir a experienciar eventos de vida positivos em comparação com os outros (Weinstein, 1980). Estes estudos clássicos apresentaram o efeito e a sua robustez recorrendo a eventos como divorciar-se, sofrer um ataque cardíaco, ser despedido (eventos negativos; Perloff & Fetzner, 1986; Weinstein, 1980), receber uma boa proposta de emprego ou viver além dos 80 anos (eventos positivos; Weinstein, 1980). Outros autores expandiram as demonstrações de otimismo irrealista a domínios diferentes. Por exemplo, mostraram que as pessoas tendem a subestimar os seus tempos de conclusão de tarefas, antecipando que vão terminar as respetivas tarefas mais cedo do que realmente acontece (Buehler, Griffin, & Ross, 1994), e a sobrestimar a duração das suas relações românticas (Epley & Dunning, 2006; MacDonald & Ross, 1999).

Sharot e colaboradores (2011) questionaram-se acerca da forma como as pessoas conseguem manter esta tendência para o otimismo irrealista em relação ao seu futuro apesar de encontrarem frequentemente informação real que desafia essas crenças enviesadas. Para examinar esta questão, os autores introduziram uma tarefa que permitiu quantificar o ajustamento das crenças de vulnerabilidade dos participantes do seu estudo em resposta a nova informação. A tarefa consiste, num primeiro momento, em pedir que os participantes estimem a sua probabilidade de vir a experienciar 80 eventos de vida adversos diferentes (por exemplo: ser traído pelo marido/esposa, morrer antes dos 80 anos, vir a sofrer de depressão, insuficiência cardíaca, obesidade, abuso de drogas ou Alzheimer, ser despedido, divorciar-se ou ser vítima de assalto), um em cada ensaio, apresentados aleatoriamente. Depois de cada ensaio, é-lhes apresentada a probabilidade média de esse evento acontecer pelo menos uma vez na vida a um indivíduo proveniente do mesmo ambiente sociocultural dos participantes. Esta probabilidade é determinada a partir de recursos *online*, não sendo incluídos eventos muito raros ou muito comuns. De seguida, numa sessão consecutiva semelhante à primeira, volta a ser pedido aos participantes que estimem a sua probabilidade de vir a experienciar todos os 80 eventos previamente expostos, um em cada ensaio. Este paradigma permite avaliar em que medida os participantes utilizaram a informação que lhes foi apresentada entre estimativas para atualizar as suas previsões. Nesse sentido, é calculada para cada evento a proporção de atualização, correspondente à diferença entre a primeira e a segunda estimativa dos participantes. A atualização absoluta média é depois comparada entre os ensaios nos quais os participantes receberam informação desejável (probabilidade de o evento negativo acontecer a um indivíduo proveniente do mesmo ambiente sociocultural mais baixa do que a primeira estimativa) e



indesejável (probabilidade de o evento negativo acontecer a um indivíduo proveniente do mesmo ambiente sociocultural mais elevada do que a primeira estimativa).

Ao contrário do esperado de acordo com algumas teorias da aprendizagem, segundo as quais as pessoas devem ajustar as suas expetativas quando confrontadas com nova informação, nomeadamente informação desconfirmatória, Sharot e colaboradores (2011) detetaram uma forte assimetria dependente da valência ao nível da atualização de crenças. Especificamente, os participantes atualizaram mais as suas crenças em resposta a informação desejável do que em resposta a informação indesejável. Por outras palavras, aprenderam mais a partir de informação que lhes ofereceu uma oportunidade de adotarem expetativas mais otimistas do que a partir de informação que contrariou a sua perspetiva positiva, sendo que o efeito foi robusto. De salientar que a atualização diferencial não foi explicada por diferenças ao nível da memória e não refletiu características específicas dos eventos adversos, como a sua familiaridade, a sua negatividade, a excitação emocional associada ou a experiência passada com os eventos, não tendo sido também resultado de diferenças ao nível das probabilidades reais dos eventos ou de um processamento diferencial de percentagens baixas e elevadas. A seletividade observada foi mediada por uma falha relativa na codificação de erros que deveriam reduzir o otimismo (Sharot et al., 2011). Os autores concluíram que o viés de aprendizagem otimista identificado quando as crenças dos participantes são desafiadas por nova informação está associado a uma falha seletiva na atualização e a uma codificação neuronal reduzida de informação indesejável acerca do futuro, podendo dizer-se, assim, que a atualização seletiva de crenças permite a manutenção do otimismo irrealista. Neste estudo, o otimismo foi definido em termos da subestimação de ocorrência de eventos negativos, mas sabe-se já que o padrão de atualização assimétrica aqui evidenciado é semelhante para eventos de vida futuros positivos e negativos (Garrett & Sharot, 2017).

O paradigma descrito foi adaptado diversas vezes e utilizado amplamente pelo grupo de investigação dos autores originais (e.g., Chowdhury, Sharot, Wolfe, Düzel, & Dolan, 2014; Korn, Sharot, Walter, Heekeren, & Dolan, 2014; Moutsiana et al., 2013; Sharot, Guitart-Masip, Chowdhury, Korn, & Dolan, 2012; Sharot, Kanai, et al., 2012). De modo mais relevante para o presente propósito, Garrett e Sharot (2014) modificaram o paradigma de modo a examinar se a atualização de crenças assimétrica previamente demonstrada se limita ou não a estimativas de risco futuro relativas ao próprio. Assim, para além de estimarem o seu risco de vir a experienciar os 80 eventos de vida adversos utilizados no estudo inicial, os participantes forneceram também estimativas das *base rates* desses eventos, isto é, da probabilidade de cada evento vir a ocorrer em média na população. Em cada ensaio, receberam depois informação

explícita acerca da frequência real dos eventos numa população demograficamente semelhante, e numa segunda sessão consecutiva voltaram a estimar ambas as probabilidades. Os autores perceberam que enquanto a integração enviesada de informação relativa ao próprio é um fenómeno robusto, mantendo-se mesmo sob diferentes definições empíricas daquilo que é considerado informação desejável, a integração enviesada de informação relativa às *base rates* da população não é tão robusta nem tão clara. Este trabalho é pertinente para o presente propósito no sentido em que aplica a tarefa de atualização de crenças de Sharot e colaboradores à análise de percepções relativas a outros, para além do próprio. Ainda assim, continua a referir-se a eventos de vida que podem vir a acontecer às pessoas no futuro, o que difere significativamente do estudo do comportamento no domínio da moralidade.

Um outro trabalho que adaptou o paradigma de Sharot e colaboradores de formas relevantes para o presente propósito foi o de Kuzmanovic, Jefferson, e Vogeley (2014). À semelhança de Garrett e Sharot (2014), estes autores também procuraram investigar se o viés otimista ao nível da atualização de crenças é ou não específico para o próprio. Assim, em metade dos 88 ensaios com eventos adversos diferentes (por exemplo: vir a sofrer de artrite, vir a ficar desempregado ou vir a ter uma infestação em casa) foram pedidas estimativas em relação ao próprio e na outra metade em relação a um outro semelhante, com a mesma idade, sexo e *background* sociocultural dos participantes. Em cada ensaio, foi pedido aos participantes que estimassem a probabilidade de os próprios ou os outros semelhantes virem a experienciar um dos eventos adversos pelo menos uma vez no futuro, foi-lhes depois apresentada a *base rate* oficial referente à probabilidade de esse evento ocorrer na população geral e subsequentemente foi-lhes pedido que fornecessem uma segunda estimativa, na qual tiveram oportunidade de rever e atualizar a estimativa inicial. Ao contrário dos estudos de Sharot e colaboradores, nos quais a primeira e a segunda estimativas eram solicitadas em duas sessões consecutivas, neste caso o pedido da primeira estimativa, a apresentação da *base rate* correspondente e o pedido da segunda estimativa foram conduzidos sucessivamente, no mesmo ensaio, de forma a minimizar possíveis efeitos de memória. Ademais, e contrariamente ao que foi comunicado aos participantes, as *base rates* foram sistematicamente manipuladas de forma a controlar e contrabalançar ao longo das condições o tamanho dos chamados erros de estimativa (isto é, as diferenças entre a primeira estimativa e a *base rate* apresentada) e, assim, controlar também a frequência de ensaios positivos/desejáveis e negativos/indesejáveis. Especificamente, as *base rates* positivas e negativas foram computadas subtraindo ou adicionando diferentes valores a partir da primeira estimativa dos participantes, independentemente do valor da estimativa. Para cada condição, os autores utilizaram para esse

cálculo os mesmos 22 valores, entre 4 e 25. Quando a *base rate* para um evento adverso era mais baixa do que a primeira estimativa do participante, essa informação foi considerada desejável ou positiva, e quando era mais elevada, foi considerada indesejável ou negativa. De modo a prevenir *base rates* negativas nos ensaios positivos ou desejáveis, as *base rates* apresentadas foram corrigidas para 1% sempre que a primeira estimativa do participante foi mais baixa do que o valor de *base rate* selecionado. De modo a prevenir *base rates* demasiado elevadas nos ensaios negativos ou indesejáveis, em particular *base rates* superiores a 100%, as *base rates* apresentadas foram corrigidas para um valor selecionado aleatoriamente entre 80% e 85% quando a soma da resposta do participante e o valor do erro excedeu 85%. Nestes estudos, as respostas foram fornecidas ajustando o número apresentado no ecrã de forma a indicar uma probabilidade absoluta entre 1% e 99% de modo completamente contínuo, sendo que no início de todas as primeiras estimativas o número no ecrã era 50%, e no início de todas as segundas estimativas foi apresentada no ecrã a primeira estimativa no respetivo ensaio. Os participantes foram encorajados a considerar a *base rate* apresentada aquando da sua segunda estimativa. Dada a transparência do *design* e o facto de as *base rates* terem sido manipuladas, os autores avaliaram cuidadosamente as possíveis suspeições dos participantes relativamente ao objetivo do estudo utilizando um método de *funnel debriefing* no final das tarefas. Os participantes foram também informados de que as *base rates* tinham sido manipuladas aleatoriamente, tendo sido explicadas as razões para este procedimento experimental.

Os resultados deste trabalho também apoiaram a existência de um viés otimista na atualização de crenças específico para o próprio, refletindo-se numa negligência seletiva de nova informação indesejável, embora os participantes com reduzido otimismo geral tenham apresentado uma atualização de crenças otimista também para os outros (Kuzmanovic et al., 2014). A relevância destes estudos para o presente propósito advém, mais uma vez, do facto de a tarefa de atualização de crenças de Sharot e colaboradores ter sido aplicada à análise de perceções relativas a outros, para além do próprio, mas também de algumas especificidades do método, nomeadamente a apresentação sucessiva do pedido da primeira estimativa, da *base rate* e do pedido da segunda estimativa no mesmo ensaio, que contribui para a transparência do *design*, e o procedimento de manipulação das *base rates*. No entanto, há que notar que quando se referiram aos outros, Kuzmanovic e colaboradores (2014) não especificaram se os participantes deviam pensar num grupo de pessoas semelhantes em termos gerais ou numa pessoa semelhante específica que conhecessem. E, além disso, o trabalho continua a debruçar-se sobre eventos de vida futuros e não sobre o comportamento no domínio moral.

Em suma, sabendo que as pessoas tendem a sofrer de um viés otimista quando efetuam previsões acerca de eventos que podem acontecer no seu futuro (otimismo irrealista; e.g., Buehler et al., 1994; Epley & Dunning, 2006; MacDonald & Ross, 1999; Perloff & Fetzer, 1986; Taylor & Brown, 1988; Weinstein, 1980, 1989), e que um dos mecanismos propostos para a manutenção deste efeito está relacionado com a forma como as pessoas atualizam as suas crenças relativas ao próprio perante informação real, ajustando as suas estimativas de forma assimétrica, isto é, mais em resposta a informação desejável do que em resposta a informação indesejável (viés de atualização otimista ou *optimistic update bias*; e.g., Chowdhury et al., 2014; Garrett & Sharot, 2014, 2017; Korn et al., 2014; Kuzmanovic et al., 2014; Moutsiana et al., 2013; Sharot, Guitart-Masip et al., 2012; Sharot, Kanai et al., 2012; Sharot et al., 2011), pretende-se investigar o fenómeno de atualização de crenças acerca do comportamento moral da humanidade enquanto possível mecanismo explicativo da manutenção do cinismo em relação ao comportamento moral das pessoas em geral. Neste caso, o foco reside no domínio do comportamento moral, ao invés dos eventos de vida futuros, e nos outros, ao invés do próprio. Para tal, é conduzido um estudo que avalia a atualização de crenças acerca de comportamentos morais e imorais dos outros por integração de informação “real”, nomeadamente *base rates* sob a forma de percentagens referentes à prevalência desses comportamentos na amostra em causa. O paradigma utilizado é, assim, baseado no original de Sharot e colaboradores, com aproximações ao método de Kuzmanovic e colaboradores (2014).

Tendo em conta o enquadramento teórico, espera-se que no caso da atualização de crenças referente ao comportamento dos outros no domínio moral se verifique um padrão assimétrico pessimista, caracterizado por maior atualização perante informação negativa ou pessimista em comparação com a atualização perante informação positiva ou otimista. Propõe-se, assim, que as pessoas não integram informação de valência oposta nas suas crenças da mesma forma, apresentando maior facilidade de assimilação de nova informação negativa ou pessimista sobre o comportamento dos outros do que de nova informação positiva ou otimista, o que impede que aprendam eficazmente que os outros não são assim *tão* maus.

Na realidade, a atribuição de maior peso à informação negativa em comparação com a informação positiva foi já constatada em alguns estudos no âmbito da atualização de crenças, sobretudo na área da economia. Em particular, alguns autores apresentaram evidências de atualização de crenças pessimista para o próprio sob determinadas circunstâncias (e.g., Coutts, 2018; Ertac, 2011). De modo mais relevante para o presente propósito, Erkal, Gangadharan, e Koh (2019) demonstraram atualização de crenças pessimista para os outros. Ainda assim, este trabalho difere do presente, não só por empregar um paradigma de atualização de crenças

significativamente distinto, mas também por focar os resultados de decisões dos outros que afetam diretamente o próprio (domínio da tomada de decisão) e não ser referente a comportamentos morais e imorais dos outros (domínio moral).

Além disto, Doosje, Spears e Koomen (1995) afirmaram que o conhecido efeito de homogeneidade do *outgroup* deve, como previamente reportado, aplicar-se a *outgroups* negativos (ou seja, quando é recebida informação negativa sobre um *outgroup*), sendo a expressão “eles são todos iguais” particularmente ilustrativa do mesmo. Os autores sugerem, no entanto, que quando os *outgroups* são positivos (ou seja, quando é recebida informação positiva sobre um *outgroup*), princípios motivacionais podem conduzir à percepção de heterogeneidade do *outgroup*, aplicando-se aqui a expressão “eles são bastante variados”. A hipótese enunciada para o presente trabalho parece alinhar-se com esta interpretação, no sentido em que quando a informação apresentada é negativa/pessimista espera-se que os outros sejam considerados homogêneos, de tal forma que aquela informação é representativa da humanidade, generalizando-se a outras amostras (“eles são todos iguais”), mas quando a informação é positiva/otimista espera-se que os outros sejam considerados heterogêneos, de tal forma que aquela informação não é representativa da humanidade, mas antes tida como uma exceção, não sendo generalizada a outras amostras (“eles são bastante variados”).

Um objetivo secundário do presente trabalho passa por avaliar o efeito BTA neste contexto e explorar uma possível relação entre o BTA e o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral. Com esse propósito, é conduzida uma segunda tarefa na qual os participantes devem indicar a sua probabilidade de desempenhar cada um dos comportamentos (morais e imorais) considerados na tarefa de atualização de crenças em comparação com as outras pessoas em geral. As respostas devem ser apresentadas numa escala de *Likert* de 9 pontos, de -4 (“É muito menos provável que eu faça isto do que as outras pessoas em geral”) a +4 (“É muito mais provável que eu faça isto do que as outras pessoas em geral”), com ponto médio em 0 (“É tão provável que eu faça isto quanto as outras pessoas em geral”). Adota-se, assim, uma medida direta do efeito BTA com foco eu-outros. A opção pela medida direta, na qual o próprio se compara com os outros na mesma escala, ao invés de uma medida indireta, na qual o próprio e os outros são avaliados em escalas separadas posteriormente comparadas, prende-se com o facto de a medida direta ser mais comumente utilizada em estudos com o efeito BTA (Alicke & Govorun, 2005) e ser potencialmente mais clara para os participantes. Ainda assim, há que ter em conta que o seu enquadramento comparativo mais forte pode fazer com que as diferenças entre o próprio e os outros sejam particularmente evidentes (Alicke & Govorun, 2005). De modo semelhante, o foco eu-outros,

no qual o próprio se compara com os outros, ao invés de outros-eu, no qual os outros são comparados com o próprio, é mais comumente adotado, apesar de também tender a conduzir logo à partida a diferenças maiores (Alicke & Govorun, 2005).

No que diz respeito a esta tarefa, espera encontrar-se, em primeiro lugar, um efeito BTA no domínio do comportamento moral. De facto, o autoenaltecimento é levado a cabo com a distorção menos óbvia e, por isso, mais dentro dos limites da realidade, quando a dimensão em causa é subjetiva ou abstrata, por oposição a objetiva ou concreta, e os critérios utilizados na avaliação do domínio moral tendem a enquadrar-se particularmente bem nesta descrição (Alicke & Govorun, 2005), ainda que sejam apresentados aos participantes comportamentos concretos. Em segundo lugar, e à semelhança dos resultados reportados por Klein e Epley (2016, 2017), que revelaram que as pessoas se sentem moralmente superiores aos outros particularmente em domínios imorais, e não tanto em domínios morais, espera-se que o efeito BTA encontrado seja mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais.

Por último, e numa perspetiva mais exploratória, testar-se-á uma possível correlação positiva entre o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral e o efeito BTA, de acordo com a qual quanto maior a atualização de crenças acerca do comportamento dos outros no domínio moral perante informação negativa ou pessimista em comparação com a atualização de crenças perante informação positiva ou otimista, maior o efeito BTA. Esta proposta baseia-se na ideia de que apesar de a tarefa de atualização de crenças ser uma tarefa na qual os participantes avaliam os outros, o modo como a abordam pode, até certo ponto, estar também relacionado com a forma como pensam sobre si próprios e como se comparam com esses outros. De facto, como evidenciado por Dunning (2001), mesmo quando o alvo não é o próprio, as pessoas tendem a fazer julgamentos que refletem favoravelmente sobre si, reforçando implícita e indiretamente a sua autoestima. Em particular, a necessidade de promover a própria autoestima pode afetar o modo como as pessoas revêem as impressões que têm dos outros (Dunning, 2001). Tendo isto em conta e sabendo que as pessoas possuem um forte desejo ou necessidade de se percecionarem como morais e éticas (Dunning, 2001) e que, tal como previamente referido, tendem a avaliar-se mais favoravelmente do que os outros também no domínio moral, sendo que esta tendência pode ser encarada como uma heurística que, entre outras funções adaptativas, contribui para a preservação de um nível relativamente elevado de autoestima (Alicke et al., 1995), é plausível que adotar e manter uma perspetiva mais pessimista em relação ao comportamento dos outros no domínio moral possa constituir um método que ajuda a perpetuar uma perspetiva otimista

relativamente ao próprio. Por outras palavras, ao constituir presumivelmente um dos mecanismos subjacentes à persistência do cinismo em relação à bondade da humanidade, a atualização de crenças acerca do comportamento dos outros no domínio moral no sentido negativo ou pessimista pode permitir uma conservação do efeito BTA e, por sua vez, uma manutenção da autoestima de cada pessoa.

Esta abordagem motivacional ao efeito encontra-se em linha com o trabalho de Ybarra (Ybarra, 1999; Ybarra & Stephan, 1996), segundo o qual as pessoas recordam preferencialmente comportamentos negativos pelos quais os outros são culpados (atribuição disposicional) e comportamentos positivos pelos quais os outros não recebem crédito (atribuição situacional), ou seja, recordam melhor os comportamentos que refletem mais negativamente sobre os outros (Ybarra & Stephan, 1996), sendo que este efeito está pelo menos parcialmente relacionado com a motivação das pessoas para melhorar a sua autoimagem moral, dado que quando a necessidade de autoenaltcimento foi removida ao fomentar a autoestima dos participantes fornecendo-lhes *feedback* positivo acerca das suas capacidades numa tarefa de raciocínio analógico, o efeito foi eliminado (Ybarra, 1996). Parece, então, que as pessoas possuem uma necessidade de percecionar os outros como moralmente maus de forma a potenciar a sua própria autoestima.

Todavia, tratando-se de correlações, não será possível constatar se a atualização assimétrica contribui para o efeito BTA e, conseqüentemente, para a manutenção da autoestima, se é o efeito BTA que motiva a atualização assimétrica em primeiro lugar, ou se ambas se verificam.

Em síntese, colocam-se as seguintes hipóteses de estudo:

H1. Espera-se verificar um padrão assimétrico pessimista ao nível da atualização de crenças referente ao comportamento dos outros no domínio moral, caracterizado por maior atualização perante informação negativa ou pessimista em comparação com a atualização perante informação positiva ou otimista.

H2.1. Espera-se encontrar um efeito BTA no domínio moral, isto é, que os participantes considerem mais provável que eles próprios desempenhem comportamentos morais e menos provável que eles próprios desempenhem comportamentos imorais do que as outras pessoas em geral.

H2.2. Espera-se que o efeito BTA encontrado seja mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais.

H3. Espera-se encontrar uma possível correlação positiva entre o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral e o efeito BTA, de

acordo com a qual quanto maior a atualização de crenças acerca do comportamento dos outros no domínio moral perante informação negativa ou pessimista em comparação com a atualização de crenças perante informação positiva ou otimista, maior o efeito BTA.

## **Método**

### **Participantes**

Foram recrutados para participação no presente estudo 120 alunos do 1º ano do Mestrado Integrado em Psicologia da Faculdade de Psicologia da Universidade de Lisboa. Após exclusão de algumas entradas não completas, a amostra de conveniência final foi constituída por 115 participantes (101 do sexo feminino) com idades compreendidas entre os 18 e os 63 anos ( $M = 20.89$ ,  $DP = 6.613$ ).

Os participantes deram o seu consentimento informado perante as condições apresentadas e receberam créditos em troca da participação voluntária no estudo.

### **Material**

Foram selecionados, a partir de um índice de comportamentos específicos no domínio moral desenvolvido por Chadwick, Bromgard, Bromgard, e Trafimow (2006), 40 comportamentos representantes de diversos construtos morais – 20 comportamentos morais (por exemplo: “Doa dinheiro para países pobres”, representante do construto “caridoso”) e 20 comportamentos imorais (por exemplo: “Foge aos impostos”, representante do construto “desonesto”). Estes 40 comportamentos foram selecionados de entre os 500 apresentados por Chadwick e colaboradores (2006) por serem representativos de diferentes construtos morais, por terem sido considerados bastante ou extremamente representativos dos respetivos construtos (média igual ou superior a 3 numa escala de *Likert* com 4 pontos) e por serem comportamentos acerca dos quais seria credível conhecer-se a *base rate* real (embora não fosse realmente conhecida). Foram excluídos comportamentos aparentemente pouco frequentes, comportamentos muito vagos, comportamentos na negativa e comportamentos muito redundantes. Alguns comportamentos sofreram ligeiras alterações, de modo a tornarem-se mais claros e/ou a estarem mais alinhados com a realidade portuguesa. A lista final de comportamentos selecionados é apresentada em anexo (ver Anexo A).



## Procedimento

O estudo foi conduzido através da plataforma *Qualtrics* nos computadores dos laboratórios da Faculdade de Psicologia da Universidade de Lisboa durante o mês de Maio de 2019. Cada participante colaborou numa única sessão com duração aproximada de 30 minutos.

Em primeiro lugar, os participantes leram um documento de consentimento informado (ver Anexo B), no qual foi dado a conhecer o objetivo e tarefa gerais do estudo, bem como alguns aspetos importantes acerca dele. Os participantes foram informados de que não existiam respostas certas e erradas, estando os investigadores apenas interessados na sua opinião sincera. Foram também assegurados de que a sua participação era voluntária e tinham o direito de desistir em qualquer altura sem prejuízo, de que as suas repostas seriam mantidas anónimas e confidenciais, estando os investigadores apenas interessados no agrupamento dos dados de todos os participantes, e de que não existiam riscos conhecidos associados à participação no estudo. Foi ainda apresentado um endereço de *e-mail* para o qual os participantes poderiam contactar caso tivessem qualquer questão e/ou caso tivessem interesse em receber, após o estudo, um relatório síntese dos principais resultados em linguagem não técnica. Considerando a informação apresentada, todos os participantes deram o seu consentimento informado para participar no estudo.

Após o consentimento informado, foi indicado aos participantes que a informação que lhes iria ser apresentada ao longo do estudo tinha sido retirada de um levantamento de atitudes e comportamentos morais de grande escala conduzido recentemente por uma equipa de investigadores de várias Universidades, no qual foram recolhidos dados de uma amostra representativa da população geral constituída por mais de 5000 indivíduos, tentando-se, desta forma, obter uma imagem fiel do mundo real, com pessoas de todos os tipos e origens da população. Seguidamente, foram apresentadas as instruções detalhadas de cada uma das 3 partes da primeira tarefa – tarefa de atualização de crenças (ver Anexo B). Como previamente referido, o paradigma utilizado corresponde a uma adaptação do original de Sharot e colaboradores (nomeadamente, Sharot et al., 2011), com aproximações ao método de Kuzmanovic e colaboradores (2014).

Na primeira parte desta tarefa (T1), foi pedido aos participantes que estimassem a percentagem de indivíduos na amostra do levantamento referido que desempenha determinado tipo de comportamento. Por exemplo, uma pergunta possível seria: “Qual é a percentagem de indivíduos nesta amostra que ajuda nas tarefas domésticas?”. Os participantes deviam escrever

o valor inteiro entre 0 e 100 correspondente à sua resposta no espaço indicado usando o teclado do computador.

Na segunda parte, foi apresentada a percentagem “real” de indivíduos na amostra do levantamento que desempenha esse mesmo comportamento. Por exemplo, para a pergunta anterior, um caso possível seria: “A percentagem real de indivíduos nesta amostra que ajuda nas tarefas domésticas é 60%.”. Cada participante visualizou 10 comportamentos morais seguidos de *feedback* otimista (isto é, *base rates* superiores à estimativa inicial do participante em T1), 10 comportamentos morais seguidos de *feedback* pessimista (isto é, *base rates* inferiores à estimativa inicial do participante em T1), 10 comportamentos imorais seguidos de *feedback* otimista (isto é, *base rates* inferiores à estimativa inicial do participante em T1) e 10 comportamentos imorais seguidos de *feedback* pessimista (isto é, *base rates* superiores à estimativa inicial do participante em T1). O mesmo participante não recebeu *feedback* otimista e pessimista para o mesmo comportamento, mas um participante pode ter recebido *feedback* otimista para determinado comportamento enquanto outro recebeu *feedback* pessimista para esse mesmo comportamento, e vice-versa. As *base rates* apresentadas foram, à semelhança de Kuzmanovic e colaboradores (2014), derivadas a partir das estimativas iniciais dos participantes, reveladas em T1. Especificamente, corresponderam a *base rates* aleatórias dentro de um intervalo de valores (entre 15% e 25%) acima e abaixo das estimativas iniciais. Quando a subtração ao valor inicial resultou num valor negativo e quando a adição resultou num valor superior a 100%, a *base rate* apresentada foi corrigida, tal como em Kuzmanovic e colaboradores (2014), no presente caso para 0% e 100%, respetivamente.

Na terceira parte (T2), foi indicado aos participantes que a referida equipa de investigadores iria repetir este estudo, conduzindo um novo levantamento de atitudes e comportamentos morais com uma amostra igualmente grande e representativa da população geral. Aqui, foi pedido que estimassem a percentagem de indivíduos na amostra deste outro levantamento que desempenha o mesmo comportamento. Por exemplo: “Pensando no novo levantamento, qual é a percentagem de indivíduos nesta amostra que ajuda nas tarefas domésticas?”. Os participantes foram informados de que podiam dar a mesma resposta que tinham dado na primeira parte ou, face à informação que receberam na segunda parte, podiam atualizar a sua resposta anterior, tendo em conta que nenhum levantamento é igual a outro, pelo que é natural que as percentagens encontradas num não sejam exatamente replicadas noutro.

Este procedimento foi repetido para cada um dos 40 comportamentos considerados, perfazendo um total de 40 ensaios. A apresentação destes ensaios foi aleatorizada.

Depois, foram apresentadas as instruções para a segunda e última tarefa – tarefa BTA (ver Anexo B). Nesta tarefa, foi pedido aos participantes que indicassem se era mais ou menos provável que eles próprios desempenhassem cada um dos 40 comportamentos previamente considerados (20 comportamentos morais e 20 comportamentos imorais) em comparação com as outras pessoas em geral. Por exemplo: “Em comparação com as outras pessoas em geral, quão provável é para si ajudar nas tarefas domésticas?”. As respostas deviam ser dadas numa escala de *Likert* de 9 pontos, de -4 a +4, sendo que -4 significava “É muito menos provável que eu faça isto do que as outras pessoas em geral”, 0 significava “É tão provável que eu faça isto quanto as outras pessoas em geral” e +4 significava “É muito mais provável que eu faça isto do que as outras pessoas em geral”. Adotou-se, assim, uma medida direta do efeito BTA (ao invés de uma medida indireta), com foco eu-outros (ao invés de outros-eu).

De seguida, foi pedido que os participantes indicassem alguns dados demográficos – sexo e idade.

Por fim, foi dado a conhecer aos participantes o objetivo do estudo de forma mais desenvolvida e específica, revelando também que as percentagens apresentadas não correspondiam, na verdade, a percentagens reais de indivíduos da amostra de um levantamento de atitudes e comportamentos morais, tendo sido criadas pelos investigadores responsáveis por este estudo. Foi ainda pedido que não divulgassem estas informações a outros potenciais participantes. O endereço de *e-mail* para esclarecimento de questões e/ou demonstração de interesse em receber, após o estudo, um relatório síntese dos principais resultados em linguagem não técnica foi lembrado e a participação foi agradecida (ver Anexo B).

Depois da aplicação, os dados recolhidos através da plataforma *Qualtrics* foram extraídos para um documento *Excel*. Após um processo de organização da base de dados, os resultados foram analisados estatisticamente com recurso ao software *Statistical Package for Social Sciences* (IBM SPSS, v.25.0).

### ***Design experimental***

A tarefa de atualização de crenças obedece a um *design* fatorial 2x2, com o *feedback* fornecido (otimista, pessimista) e a valência do comportamento (moral, imoral) como variáveis independentes intraparticipantes. A variável dependente corresponde à atualização de crenças, isto é, à diferença entre as primeiras e as segundas estimativas dos participantes.

## Resultados

### Tarefa de Atualização de Crenças

Em primeiro lugar, foram calculados os valores correspondentes à atualização de crenças dos participantes de T1 para T2, isto é, as diferenças entre as estimativas fornecidas antes e depois da apresentação da percentagem “real” de indivíduos na amostra do levantamento que desempenhava cada um dos comportamentos (i)morais em questão. No caso dos 10 comportamentos morais seguidos de *feedback* otimista e dos 10 comportamentos imorais seguidos de *feedback* pessimista, para os quais as *base rates* apresentadas foram superiores à estimativa inicial do participante, esperando-se, por isso, uma estimativa atualizada mais elevada, as diferenças foram calculadas como T2-T1. No caso dos 10 comportamentos morais seguidos de *feedback* pessimista e dos 10 comportamentos imorais seguidos de *feedback* otimista, para os quais as *base rates* apresentadas foram inferiores à estimativa inicial do participante, esperando-se, por isso, uma estimativa atualizada mais reduzida, as diferenças foram calculadas como T1-T2. Assim, em ambos os casos, diferenças positivas indicam uma mudança em direção à informação apresentada entre T1 e T2 e diferenças negativas indicam uma mudança na direção contrária à informação apresentada entre T1 e T2.

De seguida, foram calculadas as médias de todas as diferenças entre as estimativas em T1 e T2 para os casos em que o *feedback* apresentado foi otimista e o comportamento era moral (MédiaDIFOM), para os casos em que o *feedback* apresentado foi otimista e o comportamento era imoral (MédiaDIFOI), para os casos em que o *feedback* apresentado foi pessimista e o comportamento era moral (MédiaDIFPM) e para os casos em que o *feedback* apresentado foi pessimista e o comportamento era imoral (MédiaDIFPI). As estatísticas descritivas são apresentadas na Tabela 1.

**Tabela 1**

Estatísticas descritivas referentes às diferenças entre as estimativas dos participantes em T1 e T2 – Médias e Desvios-Padrão.

	MédiaDIFOM <sup>a</sup>	MédiaDIFOI <sup>b</sup>	MédiaDIFPM <sup>c</sup>	MédiaDIFPI <sup>d</sup>
M (DP)	11.42 (6.15)	10.10 (5.16)	11.17 (5.31)	11.40 (5.87)

<sup>a</sup> Média das diferenças entre as estimativas em T1 e T2 para os casos em que o *feedback* apresentado foi otimista e o comportamento era moral (fala-se aqui de *base rates* superiores aos valores estimados em T1)

<sup>b</sup> Média das diferenças entre as estimativas em T1 e T2 para os casos em que o *feedback* apresentado foi otimista e o comportamento era imoral (fala-se aqui de *base rates* inferiores aos valores estimados em T1)

<sup>c</sup> Média das diferenças entre as estimativas em T1 e T2 para os casos em que o *feedback* apresentado foi pessimista e o comportamento era moral (fala-se aqui de *base rates* inferiores aos valores estimados em T1)

<sup>d</sup> Média das diferenças entre as estimativas em T1 e T2 para os casos em que o *feedback* apresentado foi pessimista e o comportamento era imoral (fala-se aqui de *base rates* superiores aos valores estimados em T1)

De modo a testar a atualização de crenças dos participantes de T1 para T2, foi então conduzida uma ANOVA de medidas repetidas utilizando as médias das diferenças previamente calculadas para cada participante, com a valência do comportamento (moral ou imoral) e o *feedback* fornecido (otimista ou pessimista) como medidas repetidas intraparticipantes. Os resultados revelam que o efeito principal do *feedback* não é significativo,  $F(1, 114) = 1.75$ ,  $p = .188$ ,  $\eta_p^2 = .02$ , indicando que a atualização não é significativamente maior quando o *feedback* apresentado é pessimista do que quando o *feedback* apresentado é otimista. Verifica-se, porém, um efeito principal significativo da valência do comportamento,  $F(1, 114) = 5.32$ ,  $p = .023$ ,  $\eta_p^2 = .05$ , segundo o qual existe maior atualização quando os comportamentos apresentados são morais do que quando os comportamentos apresentados são imorais. Os resultados revelam ainda uma interação valência x *feedback* significativa,  $F(1, 114) = 5.58$ ,  $p = .020$ ,  $\eta_p^2 = .05$ , indicando que, na realidade, o efeito do *feedback* fornecido aos participantes entre T1 e T2 depende da valência dos comportamentos em causa. A Figura 1 apresenta uma representação gráfica da atualização de crenças em função da valência do comportamento e do *feedback* fornecido.

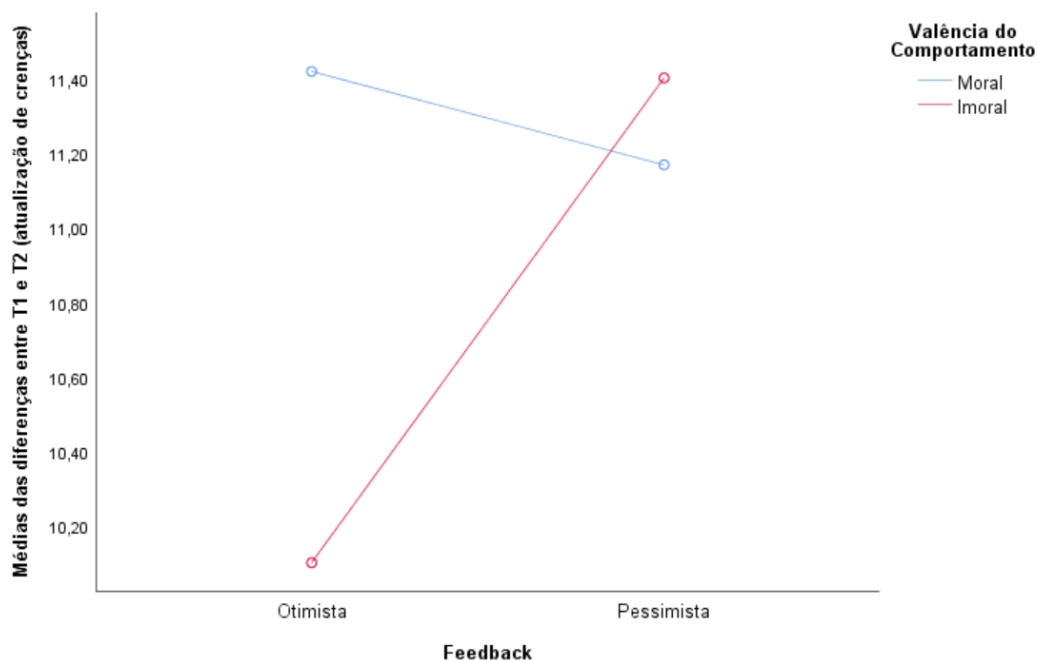


Figura 1. Atualização de crenças em função da valência do comportamento e do *feedback* fornecido.

Importa referir que esta análise recorrendo à ANOVA de medidas repetidas foi também conduzida excluindo os participantes cujas médias das diferenças entre as estimativas em T1 e T2 (MédiaDIFOM e/ou MédiaDIFOI e/ou MédiaDIFPM e/ou MédiaDIFPI) corresponderam a valores negativos, reveladores de um comportamento contrário ao esperado do ponto de vista racional na generalidade dos ensaios. O padrão de resultados não se revelou diferente. O mesmo se verificou com a exclusão dos participantes cujas médias das diferenças entre as estimativas em T1 e T2 foram iguais a 0, isto é, aqueles que tenderam a não atualizar as suas estimativas iniciais em função da nova informação. Dado que a sua inclusão não altera os resultados, estes casos foram conservados.

Numa análise adicional, recorrendo a testes t para comparar a atualização entre as diferentes condições de *feedback* (otimista vs. pessimista) no caso dos comportamentos morais e no caso dos comportamentos imorais, constata-se que a diferença na atualização é significativa quando os comportamentos em questão são imorais,  $t(114) = -2.63$ ,  $p = .010$ , IC 95% [-2.28; -0.32], de tal forma que a atualização é maior quando o *feedback* apresentado é pessimista do que quando o *feedback* apresentado é otimista, mas não é significativa quando os comportamentos são morais,  $t(114) = 0.47$ ,  $p = .640$ , IC 95% [-0.81; 1.31].

## Tarefa BTA

De forma a calcular a média relativa aos ensaios da tarefa BTA com comportamentos imorais bem como a média geral, os valores referentes aos 20 comportamentos imorais foram invertidos, de tal forma que, à semelhança dos comportamentos morais, valores positivos passam a corresponder à existência de BTA. Tanto a média geral, considerando comportamentos morais e imorais ( $M = 1.49$ ,  $DP = .642$ ), como as médias separadas, correspondentes aos ensaios com comportamentos morais ( $M = 1.16$ ,  $DP = .713$ ) e aos ensaios com comportamentos imorais ( $M = 1.82$ ,  $DP = .751$ ) permitem, efetivamente, constatar a presença de um efeito BTA, na medida em que todas diferem significativamente de zero ( $t(114) = 24.87$ ,  $p < .001$ ,  $IC\ 95\% [1.37; 1.61]$  no caso da média geral;  $t(114) = 17.43$ ,  $p < .001$ ,  $IC\ 95\% [1.03; 1.29]$  no caso da média correspondente aos ensaios com comportamentos morais;  $t(114) = 25.95$ ,  $p < .001$ ,  $IC\ 95\% [1.68; 1.96]$  no caso da média correspondente aos ensaios com comportamentos imorais). Pode dizer-se que os participantes consideraram que era mais provável que eles próprios desempenhassem comportamentos morais e menos provável que eles próprios desempenhassem comportamentos imorais do que as outras pessoas em geral.

Recorrendo a um teste  $t$  para comparar a média relativa aos ensaios da tarefa BTA com comportamentos morais ( $M = 1.16$ ,  $DP = .713$ ) e a média relativa aos ensaios da tarefa BTA com comportamentos imorais ( $M = 1.82$ ,  $DP = .751$ ) verifica-se que a diferença é significativa,  $t(114) = -10.00$ ,  $p < .001$ ,  $IC\ 95\% [-0.79; -0.53]$ . Isto é, a crença dos participantes de que a sua probabilidade de desempenhar comportamentos imorais é menor do que a probabilidade das outras pessoas em geral é mais forte do que a crença de que a sua probabilidade de desempenhar comportamentos morais é maior do que a probabilidade das outras pessoas em geral. Por outras palavras, o efeito BTA é mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais.

## Relação Entre a Tarefa de Atualização de Crenças e a Tarefa BTA

Numa fase posterior, com o propósito de explorar a relação entre a atualização de crenças e o efeito BTA, foram calculadas medidas de enviesamento geral, enviesamento no caso dos comportamentos morais e enviesamento no caso dos comportamentos imorais. Para tal, começou por calcular-se as proporções de atualização quando o *feedback* fornecido aos participantes entre T1 e T2 foi otimista no caso dos comportamentos morais [ $P(\text{AtualizaçãoOtimistaMoral})$ ], atualização quando o *feedback* fornecido aos participantes entre T1 e T2 foi pessimista no caso dos comportamentos morais

[P(AtualizaçãoPessimistaMoral)], atualização quando o *feedback* fornecido aos participantes entre T1 e T2 foi otimista no caso dos comportamentos imorais [P(AtualizaçãoOtimistaImoral)] e atualização quando o *feedback* fornecido aos participantes entre T1 e T2 foi pessimista no caso dos comportamentos imorais [P(AtualizaçãoPessimistaImoral)]. Os valores finais corresponderam à média das proporções para cada um dos respetivos comportamentos, que, por sua vez, foram obtidas através da divisão entre a atualização entre T1 e T2 (diferença entre as duas estimativas dos participantes; isto é, T2-T1 no caso dos comportamentos morais seguidos de *feedback* otimista e dos comportamentos imorais seguidos de *feedback* pessimista, e T1-T2 no caso dos comportamentos morais seguidos de *feedback* pessimista e dos comportamentos imorais seguidos de *feedback* otimista) e a diferença entre os valores de *feedback* apresentados entre T1 e T2 e a estimativa dos participantes em T1. Por exemplo:

$$P(\text{Atualização}) = \frac{T2 - T1}{|\text{Feedback} - T1|}$$

Especificamente:

$$P(\text{AtualizaçãoOtimistaMoral}) = \frac{T2 - T1}{|\text{Feedback} - T1|}$$

para todos os casos em que o *feedback* foi otimista e o comportamento moral

$$P(\text{AtualizaçãoPessimistaMoral}) = \frac{T1 - T2}{|\text{Feedback} - T1|}$$

para todos os casos em que o *feedback* foi pessimista e o comportamento moral

$$P(\text{AtualizaçãoOtimistaImoral}) = \frac{T1 - T2}{|\text{Feedback} - T1|}$$

para todos os casos em que o *feedback* foi otimista e o comportamento imoral

$$P(\text{AtualizaçãoPessimistaImoral}) = \frac{T2 - T1}{|\text{Feedback} - T1|}$$

para todos os casos em que o *feedback* foi pessimista e o comportamento imoral

Calculou-se também a proporção de atualização quando o *feedback* fornecido aos participantes entre T1 e T2 foi otimista no caso de todos os comportamentos [P(AtualizaçãoOtimista)], determinando a média das P(AtualizaçãoOtimistaMoral) e P(AtualizaçãoOtimistaImoral), e a proporção de atualização quando o *feedback* fornecido aos



participantes entre T1 e T2 foi pessimista no caso de todos os comportamentos [P(AtualizaçãoPessimista)], determinando a média das P(AtualizaçãoPessimistaMoral) e P(AtualizaçãoPessimistaImoral).

A medida de enviesamento geral foi então obtida através da seguinte fórmula:

$$\text{Enviesamento Geral} = \frac{P(\text{AtualizaçãoPessimista})}{P(\text{AtualizaçãoPessimista}) + P(\text{AtualizaçãoOtimista})}$$

E as medidas de enviesamento no caso dos comportamentos morais e no caso dos comportamentos imorais foram obtidas através das seguintes fórmulas:

$$\text{Enviesamento Moral} = \frac{P(\text{AtualizaçãoPessimistaMoral})}{P(\text{AtualizaçãoPessimistaMoral}) + P(\text{AtualizaçãoOtimistaMoral})}$$

$$\text{Enviesamento Imoral} = \frac{P(\text{AtualizaçãoPessimistaImoral})}{P(\text{AtualizaçãoPessimistaImoral}) + P(\text{AtualizaçãoOtimistaImoral})}$$

Valores de 0.5 nestas medidas indicam ausência de enviesamento, valores superiores a 0.5 indicam enviesamento pessimista e valores inferiores a 0.5 indicam enviesamento otimista.

Finalmente, de modo a testar a correlação entre as medidas de enviesamento e o efeito BTA, foi calculado o coeficiente de correlação de Pearson entre a medida de enviesamento geral e a média de todos os valores obtidos na tarefa BTA. Dado que o efeito de atualização de crenças se revelou diferente para comportamentos morais e imorais e que, de acordo com a literatura, o efeito BTA no domínio moral também difere em função da valência dos comportamentos em causa (Klein & Epley, 2016, 2017), tal como confirmado no presente trabalho, foi também calculado o coeficiente de correlação de Pearson entre a medida de enviesamento no caso dos comportamentos morais e a média dos valores obtidos na tarefa BTA referentes aos 20 comportamentos morais, e entre a medida de enviesamento no caso dos comportamentos imorais e a média dos valores obtidos na tarefa BTA referentes aos 20 comportamentos imorais. De salientar, mais uma vez, que os valores referentes aos 20 comportamentos imorais foram invertidos para o cálculo da média geral e da média relativa aos comportamentos imorais, pelo que, tal como no caso dos comportamentos morais, valores positivos correspondem à presença de efeito BTA. Nenhuma das correlações se revelou

significativa ( $r = -.02, p = .853$  no geral;  $r = -.02, p = .797$  para os comportamentos morais;  $r = .08, p = .416$  para os comportamentos imorais).

Mais uma vez, a análise foi também conduzida excluindo os participantes cujas médias das proporções de atualização ( $P(\text{AtualizaçãoOtimista})$  e/ou  $P(\text{AtualizaçãoPessimista})$  e/ou  $P(\text{AtualizaçãoOtimistaMoral})$  e/ou  $P(\text{AtualizaçãoPessimistaMoral})$  e/ou  $P(\text{AtualizaçãoOtimistaImoral})$  e/ou  $P(\text{AtualizaçãoPessimistaImoral})$ ) corresponderam a valores negativos e o padrão de resultados não se revelou diferente. O mesmo se verificou com a exclusão dos participantes cujas médias das proporções de atualização foram iguais a 0. Novamente, dado que a sua inclusão não altera os resultados, estes casos foram conservados.

## Discussão

Com o presente trabalho procurou compreender-se de que forma as crenças negativas em relação ao comportamento moral da humanidade persistem, apesar de vivermos, em grande parte, rodeados de pessoas que presumivelmente consideramos boas do ponto de vista moral e de sermos confrontados no dia-a-dia tanto com boas ações como com más ações por parte dos outros. Recorrendo a uma adaptação do paradigma de Sharot e colaboradores (nomeadamente, Sharot et al., 2011), com aproximações ao método de Kuzmanovic e colaboradores (2014), foi investigado o fenómeno de atualização assimétrica de crenças enquanto possível mecanismo explicativo da manutenção do cinismo em relação ao comportamento moral das pessoas em geral. Este estudo avaliou a atualização de crenças acerca de comportamentos morais e imorais dos outros por integração de informação “real” otimista e pessimista, nomeadamente *base rates* sob a forma de percentagens referentes à prevalência desses comportamentos na amostra em causa. De modo secundário, procurou-se também avaliar o efeito BTA no domínio moral e explorar uma possível relação entre o BTA e o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral.

Tendo em conta a literatura, era esperado verificar-se um padrão assimétrico pessimista ao nível da atualização de crenças referente ao comportamento dos outros no domínio moral, caracterizado por maior atualização perante *feedback* negativo ou pessimista em comparação com a atualização perante *feedback* positivo ou otimista (H1). Por outras palavras, esperava-se uma maior facilidade de assimilação de nova informação negativa ou pessimista sobre o comportamento dos outros do que de nova informação positiva ou otimista. Os resultados não apoiaram por completo esta hipótese. O efeito principal do *feedback* não foi significativo, isto é, a atualização não foi significativamente maior quando o *feedback* apresentado era pessimista

do que quando o *feedback* apresentado era otimista. Não obstante, para além deste resultado e da constatação de um efeito principal significativo da valência do comportamento, segundo o qual se observou maior atualização quando os comportamentos apresentados eram morais do que quando os comportamentos apresentados eram imorais, verificou-se ainda um efeito de interação significativo, indicando que, na realidade, o efeito do *feedback* fornecido aos participantes entre T1 e T2 dependeu da valência dos comportamentos em causa. Mais especificamente, percebeu-se que a diferença na atualização foi significativa quando os comportamentos em questão eram imorais, de tal forma que a atualização foi maior quando o *feedback* apresentado foi pessimista do que quando o *feedback* apresentado foi otimista, mas não foi significativa quando os comportamentos eram morais. Pode, então, dizer-se que o efeito esperado, de atualização assimétrica pessimista de crenças acerca do comportamento dos outros no domínio moral, apenas se verificou no caso dos comportamentos imorais, para os quais os participantes atualizaram significativamente mais as suas estimativas após receberem *feedback* pessimista do que após receberem *feedback* otimista. No caso dos comportamentos morais, não foram observadas diferenças significativas.

Esta assimetria nos resultados não é totalmente surpreendente. Na verdade, diversos efeitos no âmbito da Psicologia Social, tais como o autoenaltimento e os vieses comparativos, são mais evidentes no caso de dimensões negativas do que no caso de dimensões positivas. Nomeadamente, como já referido, os estudos recentes de Klein e Epley (2016, 2017) revelaram que as pessoas se sentem moralmente superiores aos outros particularmente em domínios imorais, e não tanto em domínios morais. Por outras palavras, as pessoas consideram-se mais “menos más do que os outros” do que “melhores do que os outros”, e este efeito assimétrico é robusto, emergindo em previsões comportamentais, em previsões de reações emocionais e nas inferências de caráter que são efetuadas a partir do comportamento (Klein & Epley, 2016, 2017). De forma consistente com estes resultados, Steimer e colaboradores (2019) verificaram que as diferenças eu-outros ao nível da ponderação que os participantes fizeram acerca do significado do comportamento passado no seu estudo foram mais perceptíveis no caso dos comportamentos negativos do que no caso dos comportamentos positivos.

Estes resultados também fazem sentido à luz das evidências da existência de um viés de negatividade relatadas previamente, sobretudo no contexto da formação de impressões e do julgamento moral (e.g., Fiske, 1980; Klein & O’Brien, 2016; Martijn et al., 1992; Reeder & Coover, 1986; Risky & Birnbaum, 1974; Skowronski & Carlston, 1987, 1992). Segundo estes autores, os comportamentos imorais evocam avaliações morais mais fortes do que os comportamentos morais, pelo que é plausível que até a atualização pessimista de crenças

perante nova informação relevante para a avaliação no domínio moral seja evidente no caso dos comportamentos imorais, mas não no caso dos comportamentos morais.

Remetendo para a Figura 1 apresentada na secção de Resultados, referente à atualização de crenças em função da valência do comportamento e do *feedback* fornecido, constata-se que o ponto inferior esquerdo é aquele que mais se afasta dos restantes. Isto é, destaca-se a menor atualização de crenças perante *feedback* otimista no caso dos comportamentos imorais. Assim, é possível que a manutenção do cinismo em relação ao comportamento dos outros no domínio moral esteja, na realidade, menos relacionada com uma maior assimilação de informação pessimista e mais relacionada com uma menor assimilação de informação otimista, pelo menos no caso de comportamentos imorais. De notar que em ambos os cenários se trata de uma atualização de crenças assimétrica pessimista. Assim, quando os participantes percebem que os outros desempenham um comportamento moral mais ou menos do que o inicialmente estimado ou que as pessoas desempenham um comportamento imoral mais do que o inicialmente estimado, ocorre alguma atualização, mas quando percebem que as pessoas desempenham um comportamento imoral menos do que o inicialmente estimado, ocorre significativamente menos atualização. Parece, deste modo, que os participantes estão (mais) dispostos a considerar que as pessoas são mais morais do que pensavam, menos morais, e mais imorais. Mas não estão tão dispostos a aceitar que as pessoas são menos imorais do que pensavam. A crença na imoralidade dos outros aparenta, assim, ser mais forte, mais enraizada e, deste modo, mais difícil de ser alterada, o que pode possuir importantes implicações, por exemplo, ao nível da reabilitação de criminosos na comunidade.

Esta interpretação vai ao encontro da sugestão de que as implicações de comportamentos muito negativos são mais difíceis de contradizer com base em comportamentos positivos do que as implicações de comportamentos muito positivos com base em comportamentos negativos, de tal forma que pessoas que se envolveram em comportamentos extremamente imorais parecem continuar a ser perçecionadas negativamente, mesmo que o seu restante comportamento seja exemplar (Skowronski & Carlston, 1992). De relembrar que Skowronski e Carlston (1992) afirmaram mesmo que as implicações de um ato imoral, ainda que possam ser eventualmente compensadas por benevolência suficiente e igualmente extrema, podem demorar uma vida a ser dissipadas. Também Risky e Birnbaum (1974) apontaram para esta noção de que tendo cometido ações muito imorais, as pessoas não devem esperar conseguir redimir-se com boas ações.

No que diz respeito à tarefa BTA, era esperado encontrar, em primeiro lugar, um efeito BTA no domínio do comportamento moral aqui investigado (H2.1). Os resultados apoiaram

esta hipótese: os participantes consideraram que era mais provável que eles próprios desempenhassem comportamentos morais e menos provável que eles próprios desempenhassem comportamentos imorais do que as outras pessoas em geral, tanto examinando a média geral relativa a todos os ensaios da tarefa BTA, como examinando as médias referentes aos ensaios com comportamentos morais e aos ensaios com comportamentos imorais separadamente. Como já referido, este resultado enquadra-se com a noção de que o autoenaltcimento é levado a cabo com a distorção menos óbvia e, por isso, mais dentro dos limites da realidade, quando a dimensão em causa é subjetiva ou abstrata, como deve ser o caso da dimensão moral (Alicke & Govorun, 2005).

Em segundo lugar, era esperado que o efeito BTA encontrado fosse mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais (H2.2). Mais uma vez, os resultados apoiaram a hipótese: a média referente aos ensaios da tarefa BTA com comportamentos imorais foi significativamente superior à média referente aos ensaios da tarefa BTA com comportamentos morais. Isto é, a crença dos participantes de que a sua probabilidade de desempenhar comportamentos imorais é menor do que a probabilidade das outras pessoas em geral foi mais forte do que a crença de que a sua probabilidade de desempenhar comportamentos morais é maior do que a probabilidade das outras pessoas em geral. Este resultado reproduz o padrão reportado por Klein e Epley (2016, 2017), que revelaram que as pessoas se sentem moralmente superiores aos outros particularmente em domínios imorais, e não tanto em domínios morais.

Klein e Epley (2016, 2017) sugeriram uma explicação cognitiva para o efeito, de acordo com a qual a assimetria é produzida, pelo menos em parte, pela diferença de perspetivas que as pessoas adotam quando avaliam o seu próprio comportamento (perspetiva interna, apoiada em informação como intenções, motivos e desejos, que conduz a uma avaliação do comportamento consistente com uma autoimagem positiva) e o dos outros (perspetiva externa, apoiada em informação como *base rates* da população e observações do comportamento, que conduz a previsões mais simétricas acerca da propensão dos outros para comportamentos morais e imorais). Segundo os autores, tal deve originar uma diferença eu-outros maior nas avaliações de ações relativamente imorais (“que só os outros é que desempenham, dado que as minhas têm por base intenções boas, não más”) do que nas avaliações de ações relativamente morais (“que tanto eu como os outros desempenhamos”). No entanto, não excluem uma explicação mais motivacional, admitindo a possibilidade de o efeito corresponder a um mecanismo de defesa perante informação potencialmente ameaçadora para o autoconceito. Isto é, na medida em que a maioria das pessoas acredita ser moral, imaginar-se a cometer uma ação imoral pode

ser ameaçador para a sua identidade e, por isso, desencadear processos de raciocínio motivado com o objetivo de defender um autoconceito positivo pré-existente, sendo que o mesmo não se verifica no caso de ações morais (Klein & Epley, 2016, 2017).

Independentemente da explicação para o efeito, o presente resultado é interessante em si mesmo por constituir uma reprodução do padrão reportado por Klein e Epley (2016, 2017) com uma amostra portuguesa, ou seja, numa cultura menos individualista do que a dos Estados Unidos da América (e.g., Hofstede, 1980). Tal significa que mesmo no contexto de uma mentalidade mais coletivista as pessoas continuam a autoenaltecer-se e, em particular, a acreditar que são mais “menos imorais” do que “mais morais” em comparação com os restantes membros do seu grupo.

Em último lugar, e numa perspetiva mais exploratória, foi sugerida uma possível correlação positiva entre o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral e o efeito BTA. Assim, era esperado que quanto maior fosse a atualização de crenças acerca do comportamento dos outros no domínio moral perante informação negativa ou pessimista em comparação com a atualização de crenças perante informação positiva ou otimista, maior seria o efeito BTA encontrado (H3). Os resultados não apoiaram esta hipótese, uma vez que as correlações entre a medida de enviesamento geral e a média de todos os valores obtidos na tarefa BTA, entre a medida de enviesamento no caso dos comportamentos morais e a média dos valores obtidos na tarefa BTA referentes aos comportamentos morais e entre a medida de enviesamento no caso dos comportamentos imorais e a média dos valores obtidos na tarefa BTA referentes aos comportamentos imorais não se revelaram significativas.

Esta hipótese foi originalmente avançada, sobretudo, no âmbito de uma abordagem motivacional ao efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral. Especificamente, foi sugerido que a atualização de crenças acerca do comportamento dos outros no domínio moral no sentido negativo ou pessimista poderia permitir uma conservação do efeito BTA e, por sua vez, uma manutenção da autoestima de cada pessoa, na medida em que adotar e manter uma perspetiva mais pessimista em relação ao comportamento dos outros no domínio moral pode constituir um método que ajuda a perpetuar uma perspetiva otimista relativamente ao próprio. No entanto, a ausência de correlações significativas entre o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral e o efeito BTA não apoia esta explicação motivacional, baseada numa questão de desejabilidade, ainda que continue a verificar-se a possibilidade de esta correlação existir apenas no caso de determinados comportamentos específicos.

Não se tratando de uma questão de desejabilidade, é possível que o efeito de atualização cínica de crenças acerca do comportamento dos outros no domínio moral observado no caso dos comportamentos imorais corresponda a um exemplo de viés confirmatório. Isto é, é possível que os participantes tenham ajustado mais as suas crenças em resposta a informação consistente com a sua convicção prévia de que o comportamento dos outros é maioritariamente mau do ponto de vista moral (*feedback* pessimista para comportamentos imorais) do que em resposta a informação inconsistente com essa convicção prévia (*feedback* otimista para comportamentos imorais). Embora admissível, esta possibilidade não pôde ser testada, dado que as crenças iniciais dos participantes (os chamados *priors*) não foram especificamente analisadas.

Identifica-se ainda uma última alternativa: é possível que se trate de uma questão mais adaptativa, no sentido em que o comportamento dos outros no domínio moral pode afetar o próprio diretamente, estimulando uma necessidade de proteção que conduz as pessoas a perceberem os outros como moralmente maus e a atualizar as suas crenças nesse sentido. No fundo, fala-se aqui em esperar o pior dos outros de modo a estarem preparadas para esse cenário. De facto, tal como sugerido por Ybarra (1999), enfatizar o que há de mau nos outros pode ajudar as pessoas a reforçar autoimagens positivas, mas pode, por outro lado, ser funcional de forma mais geral e tornar-se adaptativo, dado que a perceção social deve ser, pelo menos em parte, guiada por motivações de autoproteção e pela necessidade de determinar a segurança do ambiente. Esta sugestão encontra-se em linha com o facto de o efeito ter sido verificado apenas no caso dos comportamentos imorais, mais potencialmente perigosos ou ameaçadores para o próprio.

## **Implicações**

Os resultados do presente trabalho apresentam potenciais implicações e contributos tanto para a Psicologia Social como para a sociedade em geral.

No caso da Psicologia Social, este trabalho vem, em primeiro lugar, replicar as evidências da existência de um efeito semelhante ao efeito BTA geral documentado em domínios não morais ao nível dos julgamentos comparativos no domínio do comportamento moral, bem como corroborar a proposta de que as pessoas se sentem moralmente superiores aos outros particularmente em domínios imorais e não tanto em domínios morais (Klein e Epley, 2016, 2017), contribuindo, desta forma, para a literatura sobre vieses comparativos. Em segundo lugar, vem acrescentar conhecimento à literatura sobre o viés de negatividade. A

noção de que o “mau” é mais forte do que o “bom”, de tal forma que a informação negativa apresenta globalmente maior poder em comparação com a informação positiva, tem sido amplamente demonstrada por diversos autores, tal como é evidenciado nos trabalhos de revisão de Baumeister e colaboradores (2001) e Rozin e Royzman (2011), estendendo-se ao contexto da formação e revisão de impressões e do julgamento moral de outros individuais (e.g., Fiske, 1980; Klein & O’Brien, 2016; Martijn et al., 1992; Reeder & Coover, 1986; Risky & Birnbaum, 1974; Skowronski & Carlston, 1987, 1992). O presente trabalho adianta que o “mau” é mais forte do que o “bom” no domínio moral também quando são feitos julgamentos relativos a grupos de outros considerados como um todo, para além da avaliação de indivíduos particulares (pelo menos no caso de comportamentos imorais). Há que ter em conta que as previsões comportamentais referentes a indivíduos e as previsões comportamentais referentes a populações tendem a diferir, com os indivíduos específicos a serem geralmente avaliados de forma mais favorável ou moral do que os grupos (Cricher & Dunning, 2013, 2014). Ainda assim, tudo indica que o viés de negatividade no domínio da formação de impressões e do julgamento moral dos outros se verifica em ambos os casos. De resto, estes resultados contribuem para uma compreensão mais aprofundada do processo de atualização de crenças acerca dos outros no domínio moral, que se revela assimétrico, e, em particular, do fenómeno de cinismo em relação à bondade da humanidade, que parece ser predominantemente fomentado pela forma como a informação de valência oposta relativa a comportamentos imorais é diferencialmente processada.

No caso da sociedade em geral, possuir um sistema cognitivo enviesado, que aprende preferencialmente em resposta a informação moral negativa/pessimista em comparação com informação moral positiva/otimista acerca dos outros (pelo menos quando se trata de comportamentos imorais) deve originar consequências ao nível do comportamento e das decisões sociais, tanto as mais triviais e recorrentes, como decisões menos frequentes e potencialmente mais importantes. Efetivamente, e como já referido, qualquer processo que promova o cinismo relativamente aos outros ou contribua para a sua manutenção pode também conduzir a uma redução das tendências pró-sociais, nomeadamente da confiança, da cooperação e do compromisso com os pares, e, eventualmente, contribuir para a intensificação do conflito humano. Este ponto de vista encontra-se evidenciado nas observações de Epley e colaboradores (2006) a respeito do fenómeno que denominaram egoísmo reativo, relacionado com a estimulação do comportamento egoísta do próprio no seguimento da ativação de teorias acerca do comportamento provavelmente egoísta dos outros, assim como na profecia autoconfirmatória avançada por Miller e colaboradores no contexto da norma do *self-interest*



(Miller, 1999), que indica que a crença de que o *self-interest* afeta fortemente o comportamento dos outros incentiva as pessoas a comportarem-se também elas de formas egoístas. No fundo, sugere-se que as crenças em relação ao comportamento dos outros podem influenciar a forma como o próprio decide comportar-se, e que se as pessoas acreditam, em geral, que o comportamento dos outros é maioritariamente mau do ponto de vista moral, então é provável que também elas se comportem mais imoralmente em resposta. Especificamente, se as pessoas tendem a atualizar as suas crenças acerca do comportamento dos outros no domínio moral de forma assimétrica, integrando mais prontamente informação negativa/pessimista do que informação positiva/otimista, pelo menos quando se trata de comportamentos imorais, faz sentido que o seu próprio comportamento seja condizente com esse raciocínio.

À luz dos presentes resultados, espera-se, assim, que a atualização assimétrica de crenças acerca do comportamento dos outros no domínio moral, através da consequente consolidação do cinismo em relação à bondade da humanidade, exerça influência na “cascata” de comportamentos imorais desempenhados pelas pessoas, embora não tanto na “cascata” de comportamentos morais. Por exemplo, o açambarcamento de papel higiénico observado um pouco por todo o mundo no início do período de confinamento decorrente da pandemia COVID-19 no ano de 2020 pode ter sido resultado de um processo deste género: na medida em que esperavam que os outros se comportassem imoralmente, comprando quantidades excessivas de papel higiénico, e perante informação real que confirmou ou até mesmo excedeu esta expectativa negativa, perpetuada pelos meios de comunicação social, as pessoas consideraram adequado, no sentido de se prevenirem, adquirir elas próprias este bem de forma descomedida, muito para além das suas necessidades reais. Algo semelhante pode verificar-se no caso da fuga aos impostos: se é recebida a informação de que muitas empresas implementam estratégias de fuga aos impostos, então deve ser justificável que, numa tentativa de proteção, o próprio também o faça.

### **Limitações e Direções Futuras**

Uma primeira limitação do presente trabalho diz respeito ao facto de ser possível que os participantes tenham compreendido o objetivo real do estudo e detetado a respetiva variável de interesse, isto é, terem percebido que os investigadores estavam interessados na diferença entre as suas estimativas nos dois momentos em função da valência da informação apresentada e da valência dos comportamentos em causa. Kuzmanovic e colaboradores (2014), por exemplo, avaliaram cuidadosamente as possíveis suspeições dos participantes relativamente ao

objetivo do estudo utilizando um método de *funnel debriefing* no final das tarefas. No presente caso, não foi aplicado um *debriefing* detalhado, na medida em que os participantes foram informados quanto às especificidades do estudo, mas não foram questionados acerca de possíveis desconfiâncias. Realmente, as características deste estudo, à semelhança do de Kuzmanovic e colaboradores (2014), tornam-no especialmente sensível à suspeição, começando pela apresentação sucessiva do pedido da primeira estimativa, da *base rate* e do pedido da segunda estimativa no mesmo ensaio, que deve ter contribuído para a transparência do *design*, e pelo facto de as *base rates* apresentadas terem sido manipuladas. Em relação a este último ponto, tal como advertido por autores como Garrett e Sharot (2017), a invalidade de estatísticas inventadas, sem significado real, pode facilmente tornar-se aparente para os participantes, que podem também ter a noção de que não existem, na realidade, estatísticas verdadeiras referentes a tais questões. Ademais, nos casos nos quais houve necessidade de aplicar procedimentos de correção, quando a subtração dos valores entre 15% e 25% à estimativa inicial resultou num valor negativo e quando a adição resultou num valor superior a 100%, as *base rates* apresentadas foram de 0% e 100%, respetivamente, percentagens que podem ter sido consideradas irrealistamente reduzidas ou elevadas, reforçando as eventuais suspeições dos participantes. Parece, realmente, pouco credível que, por exemplo, 100% dos indivíduos da amostra em causa tivessem doado dinheiro para países pobres ou traído o/a parceiro/a. Uma possível solução passaria por utilizar ao nível do procedimento de correção o mesmo método aplicado na determinação inicial das *base rates*, isto é, ao invés de corrigir sistematicamente os valores para 0% e 100%, corrigir para percentagens dentro de um intervalo de valores. Kuzmanovic e colaboradores (2014) utilizaram uma técnica semelhante. Para além de reduzir o número de casos em que seriam apresentadas *base rates* totalmente extremas, irrealistamente reduzidas e elevadas, este método também evitaria a possibilidade de apresentação demasiado frequente dos mesmos valores corrigidos (neste caso, 0% e 100%). Independentemente do método selecionado, os procedimentos de correção podem sempre introduzir algumas irregularidades no *design* experimental, por implicarem uma alteração das *base rates* calculadas, que, desta forma, não correspondem a uma soma ou subtração de valores entre 15% e 25% à estimativa inicial, como acontece com as restantes.

Apesar de tudo isto, há que ter em conta que foram aplicadas estratégias com o objetivo de reduzir a suspeição por parte dos participantes. De modo particular: as *base rates* apresentadas foram, como já referido, *base rates* aleatórias dentro de um intervalo de valores acima e abaixo das estimativas iniciais dos participantes, e não *base rates* constantes; os comportamentos selecionados a partir do índice de Chadwick e colaboradores (2006) foram,

entre outros critérios, comportamentos acerca dos quais seria mais credível conhecer-se a *base rate* real, embora não fosse conhecida (por exemplo: doar sangue); foi indicado que a informação apresentada ao longo estudo tinha sido retirada de um levantamento de atitudes e comportamentos morais conduzido recentemente, com uma amostra grande e representativa da população geral, que permitiu obter uma imagem fiel do mundo real; foram pedidas estimativas iniciais e apresentadas *base rates* relativas a um levantamento de atitudes e comportamentos morais e pedidas segundas estimativas referentes a um outro levantamento – ainda que possa constituir uma limitação em si mesma, no sentido em que existe a possibilidade de a atualização encontrada ser apenas decorrente do facto de se tratarem de amostras diferentes em T1 e T2, esta estratégia permitiu, para além de reduzir a suspeição ao tornar a variável de interesse (diferença entre as estimativas nos dois momentos) um pouco menos evidente, ajudar a evitar que os participantes fornecessem em T2 o valor exato que lhes foi apresentado entre estimativas, uma vez que nada garante que um levantamento seja totalmente igual ao outro (“nenhum levantamento é igual a outro, pelo que é natural que as percentagens encontradas num não sejam exatamente replicadas noutro”), embora deva ser, até certo ponto, diagnóstico e, por isso, motivo para atualização (apesar de as diferenças entre as *base rates* apresentadas e as estimativas em T2 não terem sido analisadas, pelo que a possibilidade de os participantes terem simplesmente repetido algumas *base rates* se mantém). Um método de *debriefing* detalhado teria sido vantajoso, mas acredita-se que estas estratégias podem ter sido suficientes para contornar a suspeição dos participantes.

Uma outra limitação, também relacionada com as estimativas dos participantes e as *base rates* apresentadas, corresponde ao facto de terem sido verificados casos nos quais a estimativa inicial (T1) dos participantes foi igual a 0% ou a 100%, não tendo sido, de todo, possível apresentar *base rates* inferiores e superiores, respetivamente, nem mesmo recorrendo a procedimentos de correção. Para além de corresponder a um problema por si só, esta questão teve também impacto no cálculo das proporções de atualização e, consequentemente, nas medidas de enviesamento consideradas: foram identificados casos nos quais a diferença entre o valor da *base rate* apresentada e a estimativa do participante em T1 (que foi utilizada para obter as proporções de atualização que permitiram calcular as medidas de enviesamento) é igual a zero, dado que o participante apresentou em T1 uma estimativa de 0% ou de 100%, o que não permitiu a apresentação de valores inferiores ou superiores, respetivamente; isto é, teve de ser apresentado como *feedback* o valor “0” ou “100” e, por isso, a diferença correspondeu a “0 – 0 = 0” ou “100 – 100 = 0”; nestas situações, a proporção (que diz respeito à divisão entre a diferença entre as estimativas dos participantes em T1 e T2 e esta diferença)

corresponderia a uma divisão por zero, visto que a diferença é igual a zero; como tal é inexequível, nestes casos não foi possível calcular qualquer valor para a proporção de atualização. De qualquer forma, a possibilidade de fornecerem estimativas entre 0% e 100%, inclusive, permitiu que os participantes reportassem as suas crenças reais, livres de constrangimentos.

Há também que tomar em consideração que os conjuntos de comportamentos morais e imorais utilizados e até mesmo os próprios comportamentos dentro de cada conjunto podem não ser totalmente equivalentes. Na realidade, não foram controladas ou consideradas diversas características dos comportamentos que podem ter afetado significativamente os resultados. A título de exemplo: a valência dos comportamentos não foi confirmada; a sua positividade/negatividade, isto é, o quão negativos são os comportamentos imorais e o quão positivos são os comportamentos morais, não foi tida em conta; a desejabilidade dos comportamentos não foi considerada; não foi avaliada a sua relevância para o próprio, ou o seu potencial para afetar o próprio; a familiaridade dos comportamentos, a experiência prévia com os mesmos (no sentido de terem observado outras pessoas a desempenhá-los) e a respetiva controlabilidade também não foram avaliadas; não foi tida em consideração a frequência real dos comportamentos; e, por último, não foi avaliada a memória para as *base rates* apresentadas enquanto *feedback*. Relativamente ao último ponto, acredita-se que, tal como no trabalho de Kuzmanovic e colaboradores (2014), a apresentação sucessiva do pedido da primeira estimativa, da *base rate* e do pedido da segunda estimativa no mesmo ensaio deve ter minimizado eventuais efeitos de memória. Quanto aos restantes, sugere-se a realização de um pré-teste dos comportamentos ou a aplicação de um questionário numa tarefa isolada da tarefa principal de atualização de crenças que permita reconhecer especificidades como aquelas aqui referidas.

Destaca-se também como limitação o facto de não terem sido consideradas nas análises do presente estudo as diferenças entre as estimativas dos participantes em T1 e as *base rates* apresentadas enquanto *feedback* – os chamados erros de estimativa –, que podem ter afetado a atualização. Por exemplo, é possível que a atualização seja tanto maior (menor) quanto maior (menor) for a diferença entre a primeira estimativa do participante e o valor supostamente real que lhe é apresentado de seguida. Efetivamente, parece que são os erros de estimativa que conduzem a aprendizagem neste paradigma, no sentido em que predizem a atualização subsequente e que os mecanismos cerebrais envolvidos se encontram relacionados com a sua deteção (e.g., Sharot et al., 2011). Autores como Garrett e Sharot (2017) recomendam mesmo que os erros de estimativa sejam controlados. Ainda assim, apesar de os erros de estimativa

não terem sido considerados nas presentes análises, a sua extensão foi relativamente controlada, na medida em que, à semelhança do trabalho de Kuzmanovic e colaboradores (2014), as *base rates* apresentadas foram sistematicamente manipuladas, o que não se verificou, por exemplo, nos estudos de Sharot e colaboradores, nos quais foram apresentadas *base rates* verdadeiras e fixas, sem qualquer controlo.

Além dos erros de estimativa, seria importante considerar especificamente nas análises as estimativas dos participantes em T1, comumente designadas de *priors*. Na verdade, os efeitos encontrados podem estar, pelo menos em parte, relacionados com a posição inicial das crenças dos participantes. Por outras palavras, o pessimismo/otimismo de partida pode ter um impacto significativo na atualização, podendo mesmo verificar-se efeitos chão e/ou teto. Isto é, a apresentação inicial de estimativas muito reduzidas ou muito elevadas pode ter tornado particularmente difícil que os participantes reduzissem ou aumentassem essas estimativas em função do *feedback*, respetivamente. De facto, dependendo das crenças iniciais, é possível que se encontre mais espaço para atualizar num sentido do que no outro. Todavia, há que lembrar que o foco do presente trabalho diz respeito ao comportamento de atualização e não propriamente aos valores absolutos das estimativas dos participantes em T1 e T2 isoladamente.

Importa ainda referir que as diferenças observadas podem dever-se simplesmente a um impacto diferencial da redução ou aumento das percentagens, independentemente da valência. No paradigma utilizado neste estudo, uma verdadeira atualização de crenças no sentido inicialmente previsto, ou seja, mais em resposta a informação negativa ou pessimista do que em resposta a informação positiva ou otimista, implica desvios a partir da estimativa inicial em sentidos opostos para comportamentos morais e imorais: no caso dos comportamentos morais implica uma redução da estimativa inicial, enquanto no caso dos comportamentos imorais implica um aumento da estimativa inicial. No entanto, aquilo que se verificou foi uma maior atualização quando a *base rate* apresentada foi superior à primeira estimativa dos participantes, tanto no caso dos comportamentos imorais, como no caso dos comportamentos morais (embora neste último a diferença entre as condições de *feedback* não se tenha revelado significativa), apesar de estas *base rates* superiores possuírem significados diferentes para os dois tipos de comportamentos. Isto é, pode dar-se o caso de os participantes atualizarem mais as suas crenças quando recebem a informação de que a percentagem “real” é mais elevada do que quando recebem a informação de que a percentagem “real” é mais reduzida, independentemente da valência dessa informação. Uma forma de perceber se as diferenças observadas se devem, efetivamente, a um viés relacionado com a valência do *feedback* pode passar por medir ambas as caracterizações de comportamentos morais e imorais de maneira independente. Ou seja, ao

invés de ser pedido aos participantes que estimem sistematicamente a percentagem de pessoas que desempenha determinados comportamentos morais e a percentagem de pessoas que desempenha determinados comportamentos imorais, sugere-se pedir que estimem a percentagem de pessoas que desempenha comportamentos morais (neste caso, *base rates* mais elevadas correspondem a *feedback* otimista e *base rates* mais reduzidas correspondem a *feedback* pessimista), a percentagem de pessoas que não desempenha comportamentos morais (neste caso, *base rates* mais elevadas correspondem a *feedback* pessimista e *base rates* mais reduzidas correspondem a *feedback* otimista), a percentagem de pessoas que desempenha comportamentos imorais (neste caso, *base rates* mais elevadas correspondem a *feedback* pessimista e *base rates* mais reduzidas correspondem a *feedback* otimista) e a percentagem de pessoas que não desempenha comportamentos imorais (neste caso, *base rates* mais elevadas correspondem a *feedback* otimista e *base rates* mais reduzidas correspondem a *feedback* pessimista). Desta forma, para que a atualização assimétrica pessimista no caso dos comportamentos imorais identificada no presente trabalho seja apoiada, é necessário que ocorra maior atualização perante *base rates* mais elevadas quando é pedido que os participantes estimem a percentagem de pessoas que desempenha esses comportamentos imorais, como aqui verificado, mas também que a atualização seja maior perante *base rates* mais reduzidas quando é pedido que os participantes estimem a percentagem de pessoas que não desempenha os comportamentos imorais.

Por último, há que apontar a natureza correlacional deste trabalho enquanto importante limitação. Pretendeu-se investigar o fenómeno de atualização assimétrica pessimista de crenças acerca do comportamento moral dos outros enquanto possível mecanismo explicativo da manutenção do cinismo em relação ao comportamento da humanidade no domínio moral. Mas, na realidade, dada essa natureza correlacional do estudo, não é possível determinar se a atualização assimétrica pessimista de crenças (pelo menos no caso dos comportamentos imorais) constitui um mecanismo explicativo da manutenção do cinismo em relação ao comportamento dos outros no domínio moral ou se, por outro lado, corresponde a uma consequência desse cinismo.

No que diz respeito a estudos futuros, sugere-se, em primeiro lugar, a aplicação do presente paradigma com uma importante alteração: a apresentação do *feedback* entre estimativas sob a forma de supostos exemplos de indivíduos selecionados aleatoriamente da amostra do levantamento de atitudes e comportamentos morais. No dia-a-dia, não é comum que as pessoas sejam confrontadas com *base rates* referentes ao comportamento dos outros. Na prática, obtêm essa informação indiretamente, através da observação de exemplos de

comportamento individuais, como aqueles que foram mencionados no início deste trabalho: um condutor que se atravessa numa fila de trânsito, um aglomerado de pessoas que dificultam a saída do metro ou alguém que passa à frente numa fila de supermercado. Assim, na segunda parte da tarefa de atualização de crenças, após a primeira estimativa dos participantes, seria apresentado o caso real de um indivíduo da amostra. Por exemplo, para o comportamento “doar sangue”, um caso possível seria: “O Tiago foi selecionado aleatoriamente de entre as pessoas inquiridas nesta amostra. O Tiago doa sangue”. Aqui é apresentado um exemplo positivo de um comportamento moral, mas seriam também construídos exemplos negativos de comportamentos morais (“O Tiago não doa sangue”), exemplos positivos de comportamentos imorais (“O Tiago não foge aos impostos”) e exemplos negativos de comportamentos imorais (“O Tiago foge aos impostos”). Espera-se que ocorra atualização de crenças acerca do comportamento dos outros no domínio moral mesmo após a receção de informação relativa a um só indivíduo, isto é, que se verifique um efeito de generalização, e, em particular, que esta atualização de crenças seja assimétrica no sentido pessimista, pelo menos quando se trata de comportamentos imorais.

Um outro caminho potencialmente interessante passaria por avaliar o padrão de atualização de crenças acerca do comportamento do próprio no domínio moral e compará-lo com a atualização de crenças acerca do comportamento dos outros. Para tal, seria possível aplicar o presente paradigma solicitando, para além da percentagem de indivíduos na amostra de um levantamento de atitudes e comportamentos morais que desempenha determinado comportamento (i)moral, a probabilidade de o próprio participante desempenhar esses comportamentos (i)morais. Entre estimativas, seria apresentada a probabilidade média de esse comportamento (i)moral ser desempenhado por uma pessoa semelhante ao participante, proveniente do mesmo ambiente sociocultural e com uma educação e valores próximos dos seus. Este estudo assemelhar-se-ia mais aos estudos levados a cabo por Sharot e colaboradores (e.g., Sharot et al., 2011), com a diferença de que os estímulos corresponderiam a comportamentos no domínio moral e não a eventos de vida futuros. À luz das já referidas evidências da existência de um viés de atualização otimista para o próprio ao nível da previsão de eventos de vida futuros (e.g., Chowdhury et al., 2014; Garrett & Sharot, 2014; Garrett & Sharot, 2017; Korn et al., 2014; Kuzmanovic et al., 2014; Moutsiana et al., 2013; Sharot, Guitart-Masip et al., 2012; Sharot, Kanai et al., 2012; Sharot et al., 2011), e também, por exemplo, ao nível da atratividade física e da inteligência (e.g., Eil & Rao, 2011) e ao nível dos traços de personalidade (e.g., Korn, Prehn, Walter, & Heekeren, 2012), e tendo em conta o efeito BTA no domínio moral documentado na literatura e igualmente encontrado no presente

trabalho, coloca-se a hipótese de uma interação entre o alvo (próprio vs. outros) e a valência do *feedback* recebido (moral vs. imoral), segundo a qual as pessoas atualizam mais as suas crenças acerca do comportamento dos outros no domínio moral perante *feedback* negativo ou pessimista do que perante *feedback* positivo ou otimista (pelo menos no caso dos comportamentos imorais), mas atualizam mais as suas crenças acerca do seu próprio comportamento moral perante *feedback* positivo ou otimista do que perante *feedback* negativo ou pessimista.

Sugere-se ainda a exploração de possíveis variáveis moderadoras do efeito. De modo particular, pensa-se que considerações de pertença grupal podem afetar os resultados e, nesse sentido, propõe-se a aplicação do presente paradigma num estudo no qual, ao invés de serem solicitadas estimativas acerca das outras pessoas em geral, é pedido aos participantes que estimem a percentagem de indivíduos pertencentes a um *ingroup* (por exemplo: adeptos do mesmo clube de futebol) ou a um *outgroup* (por exemplo: adeptos de um clube de futebol rival) que desempenha determinados comportamentos morais e imorais. Acredita-se que a atualização cínica de crenças acerca do comportamento dos outros no domínio moral, pelo menos no caso dos comportamentos imorais, deve ser particularmente evidente quando o alvo é membro de um *outgroup* e pode ser atenuada ou até mesmo invertida quando o alvo pertence a um *ingroup*. Esta atenuação ou possível inversão pode também verificar-se caso os outros em questão correspondam a amigos, familiares ou, de modo geral, pessoas com as quais os participantes se preocupam. De facto, Kappes, Faber, Kahane, Savulescu, e Crockett (2018) demonstraram um fenómeno de otimismo vicariante no âmbito da atualização de crenças acerca de eventos de vida futuros, que surge a partir da preocupação para com os outros. Neste trabalho, os participantes exibiram um viés de aprendizagem otimista, atualizando menos as suas crenças em resposta a informação negativa do que em resposta a informação positiva, não só em relação a si mesmos, mas também quando os alvos eram amigos ou estranhos com os quais se preocupavam, nomeadamente estranhos identificáveis através de nome e fotografia, por oposição a estranhos não identificáveis, e estranhos simpáticos/morais, por oposição a estranhos antipáticos/imorais.

Outra possível variável moderadora do efeito corresponde ao envolvimento pessoal direto do participante. É plausível que a atualização cínica de crenças acerca do comportamento dos outros no domínio moral seja amplificada quando os próprios participantes são as vítimas dos comportamentos (i)morais em causa (por exemplo: ser passado à frente por alguém numa fila do supermercado), em comparação com situações nas quais apenas testemunham esses comportamentos (por exemplo: observar alguém a ser passado à frente por outra pessoa numa



fila do supermercado). Sugere-se a realização de um estudo no qual os participantes ou sofrem diretamente ou apenas testemunham uma interação positiva ou negativa com outra pessoa, e espera-se que sofrer a interação esteja associado a uma atualização de crenças acerca do comportamento dos outros no domínio moral mais pessimista do que simplesmente testemunhar essa interação.

### **Conclusão**

O presente trabalho veio complementar o conhecimento em torno da tendência humana para pensar negativamente sobre o comportamento das outras pessoas no domínio moral, ao demonstrar que a manutenção deste cinismo deve ser, pelo menos em parte, possibilitada por um processo de atualização de crenças assimétrica pessimista. Este processo é caracterizado por uma maior facilidade de assimilação de nova informação negativa ou pessimista acerca do comportamento dos outros do que de nova informação positiva ou otimista, mesmo tratando-se de estimativas numéricas que não requerem formação de impressões, e opera particularmente quando as pessoas se deparam com comportamentos imorais, por oposição a comportamentos morais. Este trabalho confirmou ainda a existência de um efeito BTA no domínio moral, mais evidente no caso dos comportamentos imorais do que no caso dos comportamentos morais, embora as correlações entre este efeito BTA e o efeito de atualização cínica de crenças não se tenham revelado significativas. Assim, conclui-se que a atualização de crenças assimétrica pessimista acerca do comportamento imoral dos outros deve constituir um dos mecanismos subjacentes à persistência do cinismo em relação à bondade da humanidade, contribuindo para que as pessoas demonstrem dificuldade em aprender que os outros não são assim *tão* maus, e que a explicação motivacional, baseada numa questão de desejabilidade, não se deve aplicar neste contexto.

## Referências

- Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49(6), 1621-1630. doi: 10.1037/0022-3514.49.6.1621
- Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In Alicke, M. D., D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85-106). New York: Psychology Press.
- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68(5), 804-825. doi: 10.1037/0022-3514.68.5.804
- Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, 7(3), 275-295. doi: 10.1521/soco.1989.7.3.275
- Armor, D. A. (1999). The illusion of objectivity: A bias in the perception of freedom from bias. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 59, 5163.
- Balcetis, E., & Dunning, D. A. (2008). A mile in moccasins: How situational experience diminishes dispositionism in social inference. *Personality and Social Psychology Bulletin*, 34(1), 102-114. doi: 10.1177/0146167207309201
- Balcetis, E., & Dunning, D. (2013). Considering the situation: Why people are better social psychologists than self-psychologists. *Self and Identity*, 12(1), 1-15. doi: 10.1080/15298868.2011.617886
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323-370. doi: 10.1037/1089-2680.5.4.323
- Bench, S. W., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the

- self and others: The role of self-discovery metaphors and the true self. *Social Cognition*, 33(3), 169-185. doi: 10.1521/soco.2015.33.3.2
- Brandimarte, L., Vosgerau, J., & Acquisti, A. (2018). Differential discounting and present impact of past information. *Journal of Experimental Psychology: General*, 147(1), 74-92. doi: 10.1037/xge0000372
- Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social cognition*, 4(4), 353-376. doi: 10.1521/soco.1986.4.4.353
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67(3), 366-381. doi: 10.1037/0022-3514.67.3.366
- Chadwick, R. A., Bromgard, G., Bromgard, I., & Trafimow, D. (2006). An index of specific behaviors in the moral domain. *Behavior Research Methods*, 38(4), 692-697. doi: 10.3758/BF03193902
- Chowdhury, R., Sharot, T., Wolfe, T., Düzal, E., & Dolan, R. J. (2014). Optimistic update bias increases in older age. *Psychological Medicine*, 44(9), 2003-2012. doi: 10.1017/S0033291713002602
- Christy, A. G., Kim, J., Vess, M., Schlegel, R. J., & Hicks, J. A. (2017). The reciprocal relationship between perceptions of moral goodness and knowledge of others' true selves. *Social Psychological and Personality Science*, 8(8), 910-917. doi: 10.1177/1948550617693061
- Coutts, A. (2018). Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics*, 22(2), 369-395. doi: 10.1007/s10683-018-9572-5
- Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology*, 47(6), 1207-1213. doi: 10.1016/j.jesp.2011.05.001

- Critcher, C. R., & Dunning, D. (2013). Predicting persons' versus a person's goodness: Behavioral forecasts diverge for individuals versus populations. *Journal of Personality and Social Psychology*, 104(1), 28-44. doi: 10.1037/a0030836
- Critcher, C. R., & Dunning, D. (2014). Thinking about others versus another: Three reasons judgments about collectives and individuals differ. *Social and Personality Psychology Compass*, 8(12), 687-698. doi: 10.1111/spc3.12142
- Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly*, 47(1), 1-15. doi: 10.1086/268763
- Diekmann, K. A., Samuels, S. M., Ross, L., & Bazerman, M. H. (1997). Self-interest and fairness in problems of resource allocation: Allocators versus recipients. *Journal of Personality and Social Psychology*, 72(5), 1061-1074. doi: 10.1037/0022-3514.72.5.1061
- Doosje, B., Spears, R., & Koomen, W. (1995). When bad isn't all bad: Strategic use of sample information in generalization and stereotyping. *Journal of Personality and Social Psychology*, 69(4), 642-655. doi: 10.1037/0022-3514.69.4.642
- Douglas, K. M., & Sutton, R. M. (2004). Right about others, wrong about ourselves? Actual and perceived self-other differences in resistance to persuasion. *British Journal of Social Psychology*, 43(4), 585-603. doi: 10.1348/0144666042565416
- Dunning, D. (2001). On the motives underlying social cognition. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 348-374). New York: Blackwell Publishers Ltd. doi: 10.1002/9780470998519.ch16
- Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, 57(6), 1082-1090. doi: 10.1037/0022-3514.57.6.1082

- Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2), 114-138. doi: 10.1257/mic.3.2.114
- Epley, N., Caruso, E. M., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology*, 91(5), 872-889. doi: 10.1037/0022-3514.91.5.872
- Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self-or social prediction?. *Journal of Personality and Social Psychology*, 79(6), 861-875. doi: 10.1037/0022-3514.79.6.861
- Epley, N., & Dunning, D. (2006). The mixed blessings of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin*, 32(5), 641-655. doi: 10.1177/0146167205284007
- Epley, N., & Waytz, A. (2010). Mind perception. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (5th ed., pp. 498–541). doi:10.1002/9780470561119
- Erkal, N., Gangadharan, L., & Koh, B. H. (2019). By chance or by choice? Biased attribution of others' outcomes. doi: 10.2139/ssrn.3251105
- Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization*, 80(3), 532-545. doi: 10.1016/j.jebo.2011.05.012
- Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology*, 70(6), 1164-1184. doi: 10.1037/0022-3514.70.6.1164
- Fetchenhauer, D., & Dunning, D. (2006). Perceptions of prosociality and solidarity in self and others. In Fetchenhauer, D., A. Flache, B. Buunk, & S. Lindenberg (Eds.), *Solidarity*

*and prosocial behavior* (pp. 61–74). New York, NY: Springer.

- Fetchenhauer, D., & Dunning, D. (2009). Do people trust too much or too little?. *Journal of Economic Psychology*, 30(3), 263-276. doi: 10.1016/j.joep.2008.04.006
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38(6), 889-906. doi: 10.1037/0022-3514.38.6.889
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77-83. doi: 10.1016/j.tics.2006.11.005
- Garrett, N., & Sharot, T. (2014). How robust is the optimistic update bias for estimating self-risk and population base rates?. *PLoS One*, 9(6), 1-8. doi: 10.1371/journal.pone.0098848
- Garrett, N., & Sharot, T. (2017). Optimistic update bias holds firm: Three tests of robustness following Shah et al. *Consciousness and Cognition*, 50, 12-22. doi: 10.1016/j.concog.2016.10.013
- Goethals, G. R. (1986). Fabricating and ignoring social reality: Self-serving estimates of consensus. In J. M. Olson, C. P. Herman, & M. P. Zanna (Eds.), *Social comparison and relative deprivation: The Ontario symposium* (Vol. 4, pp. 135-157). Hillsdale, NJ: Erlbaum.
- Goethals, G. R., Messick, D. M., & Allison, S. T. (1991). The uniqueness bias: Studies of constructive social comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison research: Contemporary theory and research* (pp. 149–176). Hillsdale, NJ: Lawrence Erlbaum.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148-168. doi: 10.1037/a0034726

- Ham, J., & Vonk, R. (2011). Impressions of impression management: Evidence of spontaneous suspicion of ulterior motivation. *Journal of Experimental Social Psychology*, 47(2), 466-471. doi: 10.1016/j.jesp.2010.12.008
- Hartley, A. G., Furr, R. M., Helzer, E. G., Jayawickreme, E., Velasquez, K. R., & Fleenon, W. (2016). Morality's centrality to liking, respecting, and understanding others. *Social Psychological and Personality Science*, 7(7), 648-657. doi: 10.1177/1948550616655359
- Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology*, 89(6), 937-950. doi: 10.1037/0022-3514.89.6.937
- Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational behavior and human decision processes*, 78(1), 25-62. doi: 10.1006/obhd.1999.2826
- Helzer, E. G., & Dunning, D. (2012). Why and when peer prediction is superior to self-prediction: The weight given to future aspiration versus past achievement. *Journal of Personality and Social Psychology*, 103(1), 38-53.
- Heyman, G., Barner, D., Heumann, J., & Schenck, L. (2014). Children's sensitivity to ulterior motives when evaluating prosocial behavior. *Cognitive Science*, 38(4), 683-700. doi: 10.1111/cogs.12089
- Hofstede, G. (1980). Motivation, leadership, and organization: Do American theories apply abroad?. *Organizational Dynamics*, 9(1), 42-63. doi: 10.1016/0090-2616(80)90013-3
- Innes, J. M., & Zeitz, H. (1988). The public's view of the impact of the mass media: A test of the 'third person' effect. *European Journal of Social Psychology*, 18(5), 457-463. doi: 10.1002/ejsp.2420180507
- Judd, C. M., Park, B., Yzerbyt, V., Gordijn, E. H., & Muller, D. (2005). Attributions of

- intergroup bias and outgroup homogeneity to ingroup and outgroup others. *European Journal of Social Psychology*, 35(6), 677-704. doi: 10.1002/ejsp.281
- Kappes, A., Faber, N. S., Kahane, G., Savulescu, J., & Crockett, M. J. (2018). Concern for others leads to vicarious optimism. *Psychological Science*, 29(3), 379-389. doi: 10.1177/0956797617737129
- Kim, A. (2014). The curious case of self-interest: Inconsistent effects and ambivalence toward a widely accepted construct. *Journal for the Theory of Social Behaviour*, 44(1), 99-122. doi: 10.1111/jtsb.12032
- Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology*, 110(5), 660-674. doi: 10.1037/pspa0000050
- Klein, N., & Epley, N. (2017). Less evil than you: Bounded self-righteousness in character inferences, emotional reactions, and behavioral extremes. *Personality and Social Psychology Bulletin*, 43(8), 1202-1212. doi: 10.1177/0146167217711918
- Klein, N., & O'Brien, E. (2016). The tipping point of moral change: When do good and bad acts make good and bad actors?. *Social Cognition*, 34(2), 149-166. doi: 10.1521/soco.2016.34.2.149
- Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience*, 32(47), 16832-16844. doi: 10.1523/JNEUROSCI.3016-12.2012
- Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, 44(3), 579-592. doi: 10.1017/S0033291713001074
- Kruger, J., & Gilovich, T. (1999). "Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology*, 76(5), 743-753. doi: 10.1037/0022-3514.76.5.743



- Kruger, J., & Gilovich, T. (2004). Actions, intentions, and self-assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin*, 30(3), 328-339. doi: 10.1177/0146167203259932
- Kuzmanovic, B., Jefferson, A., & Vogeley, K. (2014). Self-specific optimism bias in belief updating is associated with high trait optimism. *Journal of Behavioral Decision Making*, 28(3), 281-293. doi: 10.1002/bdm.1849
- Larwood, L. (1978). Swine flu: A field study of self-serving biases. *Journal of Applied Social Psychology*, 8(3), 283-289. doi: 10.1111/j.1559-1816.1978.tb00783.x
- Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology*, 62(2), 194-198. doi: 10.1037/0021-9010.62.2.194
- Liebrand, W. B., Messick, D. M., & Wolters, F. J. (1986). Why we are fairer than others: A cross-cultural replication and extension. *Journal of Experimental Social Psychology*, 22(6), 590-604. doi: 10.1016/0022-1031(86)90052-1
- Lupfer, M. B., Weeks, M., & Dupuis, S. (2000). How pervasive is the negativity bias in judgments based on character appraisal?. *Personality and Social Psychology Bulletin*, 26(11), 1353-1366. doi: 10.1177/0146167200263004
- MacDonald, T. K., & Ross, M. (1999). Assessing the accuracy of predictions about dating relationships: How and why do lovers' predictions differ from those made by observers?. *Personality and Social Psychology Bulletin*, 25(11), 1417-1429. doi: 10.1177/0146167299259007
- Martijn, C., Spears, R., Van der Pligt, J., & Jakobs, E. (1992). Negativity and positivity effects in person perception and inference: Ability versus morality. *European Journal of Social Psychology*, 22(5), 453-463. doi: 10.1002/ejsp.2420220504

- Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology*, 21(5), 480-500. doi: 10.1016/0022-1031(85)90031-9
- Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, 54(12), 1053-1060. doi: 10.1037/0003-066X.54.12.1053
- Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology*, 74(1), 53-62. doi: 10.1037/0022-3514.74.1.53
- Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S. J., & Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences*, 110(41), 16396-16401. doi: 10.1073/pnas.1305631110
- Newman, G. E., Bloom, P., & Knobe, J. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, 40(2), 203-216. doi: 10.1177/0146167213508791
- Newman, G. E., De Freitas, J., & Knobe, J. (2015). Beliefs about the true self explain asymmetries based on moral judgment. *Cognitive Science*, 39(1), 96-125. doi: 10.1111/cogs.12134
- O'Gorman, H. J. (1975). Pluralistic ignorance and white estimates of white support for racial segregation. *Public Opinion Quarterly*, 39(3), 313-330. doi: 10.1086/268231
- O'Gorman, H. J., & Garry, S. L. (1976). Pluralistic ignorance – A replication and extension. *Public Opinion Quarterly*, 40(4), 449-458. doi: 10.1086/268331
- Perloff, L. S., & Fetzer, B. K. (1986). Self–other judgments and perceived vulnerability to victimization. *Journal of Personality and Social Psychology*, 50(3), 502-510. doi: 10.1037/0022-3514.50.3.502
- Preston, C. E., & Harris, S. (1965). Psychology of drivers in traffic accidents. *Journal of Applied Psychology*, 49(4), 284-288. doi: 10.1037/h0022453

- Pronin, E., Berger, J., & Molouki, S. (2007). Alone in a crowd of sheep: Asymmetric perceptions of conformity and their roots in an introspection illusion. *Journal of Personality and Social Psychology*, 92(4), 585-595. doi: 10.1037/0022-3514.92.4.585
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological review*, 111(3), 781-799. doi: 10.1037/0033-295X.111.3.781
- Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, 43(4), 565-578. doi: 10.1016/j.jesp.2006.05.011
- Pronin, E., & Kugler, M. B. (2010). People believe they have more free will than others. *Proceedings of the National Academy of Sciences*, 107(52), 22469-22474. doi: 10.1073/pnas.1012046108
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369-381. doi: 10.1177/0146167202286008
- Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology*, 81(1), 5-16. doi: 10.1037/0022-3514.81.1.5
- Reeder, G. D., & Coover, M. D. (1986). Revising an impression of morality. *Social Cognition*, 4(1), 1-17. doi: 10.1521/soco.1986.4.1.1
- Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights don't make up for a wrong. *Journal of Experimental Psychology*, 103(1), 171-173. doi: 10.1037/h0036892
- Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, 37(3), 322-336. doi: 10.1037/0022-3514.37.3.322

- Ross, M., & Wilson, A. E. (2002). It feels like yesterday: Self-esteem, valence of personal past experiences, and judgments of subjective distance. *Journal of Personality and Social Psychology*, 82(5), 792-803. doi: 10.1037/0022-3514.82.5.792
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296-320. doi: 10.1207/S15327957PSPR0504\_2
- Schroeder, J., Waytz, A., & Epley, N. (2017). Endorsing help for others that you oppose for yourself: Mind perception alters the perceived effectiveness of paternalism. *Journal of Experimental Psychology: General*, 146(8), 1106-1125. doi: 10.1037/xge0000320
- Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology*, 22(16), 1477-1481. doi: 10.1016/j.cub.2012.05.053
- Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, 109(42), 17058-17062. doi: 10.1073/pnas.1205828109
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, 14(11), 1475-1479. doi: 10.1038/nn.2949
- Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52(4), 689-699. doi: 10.1037/0022-3514.52.4.689
- Skowronski, J. J., & Carlston, D. E. (1992). Caught in the act: When impressions based on highly diagnostic behaviours are resistant to contradiction. *European Journal of Social Psychology*, 22(5), 435-452. doi: 10.1002/ejsp.2420220503
- Steimer, A., Mata, A., & Simão, C. (2019). Ascribing meaning to the past: Self-other differences in weighing good and bad deeds. *Social Cognition*, 37(2), 174-196. doi:

- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers?. *Acta Psychologica*, 47(2), 143-148. doi: 10.1016/0001-6918(81)90005-6
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193-210.
- van Lange, P. A. (1991). Being better but not smarter than others: The Muhammad Ali effect at work in interpersonal situations. *Personality and Social Psychology Bulletin*, 17(6), 689-693. doi: 10.1177/0146167291176012
- van Lange, P. A., & Sedikides, C. (1998). Being more honest but not necessarily more intelligent than others: Generality and explanations for the Muhammad Ali effect. *European Journal of Social Psychology*, 28(4), 675-680. doi: 10.1002/(SICI)1099-0992(199807/08)28:4<675::AID-EJSP883>3.0.CO;2-5
- Waytz, A., Schroeder, J., & Epley, N. (2014). The Lesser Minds Problem. In P. Bain, J. Vaes, & J. P. Leyens (Eds.), *Humanness and dehumanization* (pp. 49–67). London, England: Psychology Press.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5), 806-820. doi: 10.1037/0022-3514.39.5.806
- Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science*, 246, 1232-1233.
- Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, 24(12), 1251-1263. doi: 10.1177/01461672982412001
- Ybarra, O. (1999). Misanthropic person memory when the need to self-enhance is absent. *Personality and Social Psychology Bulletin*, 25(2), 261-269. doi: 10.1177/0146167299025002011

Ybarra, O., & Stephan, W. G. (1996). Misanthropic person memory. *Journal of Personality and Social Psychology*, 70(4), 691-700. doi: 10.1037/0022-3514.70.4.691

## Anexo A – Lista Final de Comportamentos Selecionados

**Tabela A1**

Lista final de comportamentos selecionados.

Construto Moral	Valência	Estímulo Original (Chadwick, Bromgard, Bromgard, & Trafimow, 2006)	Estímulo Traduzido Qual é a percentagem de indivíduos nesta amostra que...?
Caridoso	Moral	Donating money to poor countries	Doa dinheiro para países pobres
Caridoso	Moral	Volunteering at a shelter or soup kitchen	É voluntária nas campanhas do Banco Alimentar Contra a Fome [Alterado]
Caridoso	Moral	Donating blood	Doa sangue
Caridoso	Moral	Buying a homeless person lunch	Compra comida para pessoas sem abrigo
Caridoso	Moral	Lending a hand to a person with a flat tire	Auxilia uma pessoa com um pneu furado
Caridoso	Moral	Giving clothes and items no longer needed to the less fortunate	Dá roupas e outros itens de que já não necessita aos menos afortunados
Cooperativo	Moral	Pulling over when an emergency vehicle is trying to pass you	Encosta o carro à berma para deixar passar um veículo de emergência [Alterado]
Cooperativo	Moral	Cleaning up after one's self after eating so the next guy has a pleasant experience	Deixa tudo limpo depois de usar a casa de banho por consideração pela pessoa que vem a seguir [Alterado]
Amigável	Moral	Helping a person at the store when they are short on money	Ajuda uma pessoa que está numa loja e não tem dinheiro suficiente para comprar o que precisa [Alterado]
Amigável	Moral	Helping a stranger pick up stuff they have dropped	Ajuda um estranho a apanhar do chão as coisas que ele deixou cair
Amigável	Moral	Helping an elderly person cross the street	Ajuda uma pessoa idosa a atravessar a rua
Amigável	Moral	Cheering up a person who seems to be having a bad day	Anima uma pessoa que parece estar a ter um mau dia
Amigável	Moral	Helping someone find a lost dog	Ajuda alguém a encontrar um cão perdido
Honesto	Moral	Returning a lost wallet or purse with the money intact	Devolve uma carteira perdida com o dinheiro intacto
Honesto	Moral	Admitting when you are to blame when involved in a car accident	Admite quando é culpada se se envolver num acidente de carro [Alterado]

(Continua)

Construto Moral	Valência	Estímulo Original (Chadwick, Bromgard, Bromgard & Trafimow, 2006)	Estímulo Traduzido Qual é a percentagem de indivíduos nesta amostra que...?
Honesto	Moral	Telling someone (sales rep, business owner) that you accidentally broke or damaged something in the store	Avisa um empregado quando parte ou danifica um produto numa loja [Alterado]
Honesto	Moral	Correcting a cashier when they give you too much change	Avisa o empregado na caixa de uma loja quando recebe troco a mais [Alterado]
Honesto	Moral	Telling a friend your real opinion when they ask you for it	Dá a sua opinião verdadeira a um amigo quando ele a pede
Honesto	Moral	Telling the truth whenever you go to a job interview	Diz a verdade quando vai a uma entrevista de emprego
Leal	Moral	Standing up for a friend when they are being teased or harassed	Defende um amigo quando este está a ser gozado ou assediado
Não Caridoso	Imoral	Refusing aid to a parent when they are in need	Recusa dar ajuda a um dos pais quando ele precisa
Não Caridoso	Imoral	Throwing away items that could be donated	Deita fora itens que podiam ser doados
Não Caridoso	Imoral	Ignoring a homeless mother and child	Ignora uma mãe e filho sem abrigo
Não Cooperativo	Imoral	Interrupting another in a conversation because their opinions oppose your own	Interrompe uma pessoa numa conversa porque as opiniões dessa pessoa são opostas às suas
Não Cooperativo	Imoral	Leaving in the middle of a discussion	Abandona uma discussão a meio
Não Amigável	Imoral	Being rude to foreigners	É rude com estrangeiros
Não Amigável	Imoral	Making fun of someone	Goza com pessoas mais fracas [Alterado]
Não Amigável	Imoral	Talking above someone	Fala enquanto outra pessoa está a falar
Desonesto	Imoral	Cheating on wife/husband/girlfriend/boyfriend	Trai o seu parceiro romântico [Alterado]
Desonesto	Imoral	Forging a check	Foge aos impostos [Alterado]
Desonesto	Imoral	Shoplifting	Rouba acessórios ou outros itens pequenos em lojas [Alterado]
Desonesto	Imoral	Hitting a parked car and then taking off	Foge do local se bater num carro estacionado [Alterado]

(Continua)



Construto Moral	Valência	Estímulo Original (Chadwick, Bromgard, Bromgard & Trafimow, 2006)	Estímulo Traduzido Qual é a percentagem de indivíduos nesta amostra que...?
Desonesto	Imoral	Taking credit for someone else's work	Toma crédito indevido pelo trabalho de outra pessoa [Alterado]
Desonesto	Imoral	Cheating on an exam	Copia num exame
Desleal	Imoral	Having sex with a friend's significant other	Tem relações sexuais com o companheiro ou companheira de um amigo
Desleal	Imoral	Turning on a friend when they are in trouble	Se vira contra um amigo quando ele está com problemas
Desleal	Imoral	Bad mouthing a good friend behind their back	Fala mal de um bom amigo nas suas costas
Desleal	Imoral	Ditching a friend because other people think he/she is weird, etc.	Abandona um amigo porque outras pessoas o consideram esquisito
Desleal	Imoral	Making a promise and not keeping it	Faz promessas que sabe que não vai conseguir cumprir [Alterado]
Desleal	Imoral	If you tell a secret someone told you	Conta segredos que outra pessoa lhe contou

## Anexo B – Consentimento Informado e Instruções

Caro(a) participante,

Desde já, agradecemos a sua colaboração.

O presente estudo tem como objetivo perceber aquilo que as pessoas pensam acerca da natureza humana. A sua tarefa consiste em ler com atenção a informação que lhe é apresentada e indicar os valores que lhe vão ser pedidos. Não existem respostas certas ou erradas. Estamos apenas interessados na sua opinião sincera.

Este estudo é conduzido pela aluna Ana Martins, com supervisão do professor doutor André Mata, no âmbito do projeto de Mestrado em Cognição Social Aplicada da Faculdade de Psicologia da Universidade de Lisboa. O estudo será constituído por uma única sessão com duração aproximada de 30 minutos. Para participar, deverá ter 18 anos ou mais. A sua participação é totalmente voluntária e tem o direito de desistir em qualquer altura sem prejuízo. As suas repostas serão mantidas anónimas e confidenciais, sendo que não estamos interessados nas respostas individuais, mas sim no agrupamento dos dados de todos os participantes. Não existem riscos conhecidos associados à participação no estudo.

Para qualquer questão e/ou caso tenha interesse em receber, após o estudo, um relatório síntese dos principais resultados em linguagem não técnica, poderá contactar a aluna Ana Martins através do seguinte e-mail:

ana-martins3@campus.ul.pt

Considerando a informação apresentada, indique se aceita ou não participar neste estudo:

- Dou o meu consentimento para participar neste estudo.
- Não dou o meu consentimento para participar neste estudo.

--- MUDAR DE ECRÃ --- PARA QUEM DER O SEU CONSENTIMENTO ---

Muito obrigado pela sua participação!

Como apresentado inicialmente, o presente estudo tem como objetivo perceber aquilo que as pessoas pensam acerca da natureza humana. **De seguida, iremos descrever em detalhe as instruções para este estudo, sendo muito importante que as leia na totalidade e com atenção.**

A informação que lhe vai ser apresentada foi retirada de um levantamento de atitudes e comportamentos morais de grande escala conduzido recentemente por uma equipa de investigadores de várias Universidades. Neste levantamento, foi recolhida informação de uma amostra representativa da população geral constituída por mais de 5000 indivíduos. Com esta amostra tentou obter-se uma imagem fiel do mundo real, com pessoas de todos os tipos e origens da população. Dada a natureza sensível de alguns dos tópicos abordados, foi garantido aos participantes o total anonimato dos seus dados.

A sua primeira tarefa vai ser constituída por 3 partes sucessivas.

Na primeira parte, vai ser pedido que estime a percentagem de indivíduos nesta amostra que desempenha determinado tipo de comportamento. Por exemplo, uma pergunta possível seria: “Qual é a percentagem de indivíduos nesta amostra que ajuda nas tarefas domésticas?”. Não existem respostas certas ou erradas. Estamos apenas interessados na sua opinião sincera. Pode escolher qualquer valor entre 0% e 100%. Deverá escrever o valor que corresponde à sua resposta no espaço indicado usando o teclado do computador.

Na segunda parte, após a sua primeira estimativa, ser-lhe-á apresentada a percentagem real de indivíduos nesta amostra que desempenha esse mesmo comportamento. Por exemplo, para a pergunta anterior, um caso possível seria: “A percentagem real de indivíduos nesta amostra que ajuda nas tarefas domésticas é 60%.”.

Em breve, a referida equipa de investigadores vai repetir este estudo. Ou seja, vai conduzir um novo levantamento de atitudes e comportamentos morais com uma amostra igualmente grande e representativa da população geral. Na terceira parte, vai-lhe ser pedido que estime a percentagem de indivíduos na amostra deste outro levantamento que desempenha determinado tipo de comportamento. Por exemplo: “Pensando no novo levantamento, qual é a percentagem de indivíduos nesta amostra que ajuda nas tarefas domésticas?”. Pode dar a mesma resposta que deu na primeira parte ou, face à informação que recebeu na segunda parte, pode atualizar a sua resposta anterior. Tenha em conta que nenhum levantamento é igual a outro, pelo que é natural que as percentagens encontradas num não sejam exatamente replicadas noutro. Mais uma vez, não existem respostas certas ou erradas. Estamos apenas interessados na sua opinião sincera.

Este procedimento irá repetir-se para cada comportamento considerado.

--- MUDAR DE ECRÃ ---

Qual é a percentagem de indivíduos nesta amostra que X?

--- MUDAR DE ECRÃ ---

A percentagem real de indivíduos nesta amostra que X é: Y%

--- MUDAR DE ECRÃ ---

Pensando no novo levantamento, qual é a percentagem de indivíduos nesta amostra que X?

[Repetir este procedimento para cada um dos comportamentos considerados]

--- MUDAR DE ECRÃ ---

Agora vamos pedir-lhe que volte a pensar sobre todos os comportamentos apresentados anteriormente.

A sua tarefa consiste em indicar se é mais ou menos provável que você desempenhe cada um desses comportamentos em comparação com as outras pessoas em geral (não só os participantes das amostras anteriores, mas as pessoas da população em geral). Por exemplo: “Em comparação com as outras pessoas em geral, quão provável é para si ajudar nas tarefas domésticas?”

Deverá responder numa escala de -4 a +4, sendo que -4 significa “É muito menos provável que eu faça isto do que as outras pessoas em geral”, 0 significa “É tão provável que eu faça isto quanto as outras pessoas em geral” e +4 significa “É muito mais provável que eu faça isto do que as outras pessoas em geral”.

--- MUDAR DE ECRÃ ---

Em comparação com as outras pessoas em geral, quão provável é para si X?

[Repetir este procedimento para cada um dos comportamentos considerados]

--- MUDAR DE ECRÃ ---

Para terminar, pedimos-lhe que indique alguns dados demográficos:

Sexo (Masculino/Feminino)

Idade

--- MUDAR DE ECRÃ ---

Terminou o estudo.

Como lhe foi dito inicialmente, pretendemos com este estudo perceber aquilo que as pessoas pensam acerca da natureza humana. No entanto, queremos informá-lo de que o objetivo principal consiste, mais especificamente, em avaliar o processo de atualização das crenças que as pessoas têm sobre o comportamento moral das outras pessoas. Nomeadamente, procuramos perceber se essa atualização de crenças (neste estudo, a diferença entre a primeira e a segunda estimativa que fez para cada comportamento) difere em função de a informação apresentada sob a forma de percentagens ser positiva ou negativa.

As percentagens apresentadas não correspondem, na verdade, a percentagens reais de indivíduos da amostra de um levantamento de atitudes e comportamentos morais. Foram criadas pelos investigadores responsáveis pela presente investigação.

Estas informações não foram reveladas inicialmente por colocarem em causa a legitimidade dos resultados, sendo que o estudo não podia, então, ter sido conduzido de outra forma.

**Pedimos que não divulgue estas informações a outros potenciais participantes. O sucesso desta investigação científica e a validade dos dados recolhidos depende disto. Como certamente compreenderá, os próximos participantes devem responder honestamente sem ser influenciados por expectativas acerca do estudo.**

Relembramos ainda que, para qualquer questão e/ou caso tenha interesse em receber, após o estudo, um relatório síntese dos principais resultados em linguagem não técnica, poderá contactar a aluna Ana Martins através do seguinte e-mail:

ana-martins3@campus.ul.pt

Mais uma vez, muito obrigado pela sua colaboração!